

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 27, 2004, 12:01:11 ; Search time 52 Seconds  
(without alignments)  
130.407 Million cell updates/sec

Title: US-09-784-553C-19\_COPY\_29\_52  
Perfect score: 129  
Sequence: 1 FRQPDVAVKLGPDYHKIKQPMW 24

Scoring table: BLOSUM62  
Gapop 10.0 , Capext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database : A\_Geneseq\_29Jan04:\*
- 1: Geneseqp1980s:\*
  - 2: Geneseqp1990s:\*
  - 3: Geneseqp2000s:\*
  - 4: Geneseqp2001s:\*
  - 5: Geneseqp2002s:\*
  - 6: Geneseqp2003as:\*
  - 7: Geneseqp2003bs:\*
  - 8: Geneseqp2004s:\*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	129	100.0	303	4 AAB95802	Aab95802 Human pro
2	129	100.0	754	2 AAY07027	Aay07027 Breast ca
3	129	100.0	801	7 ADC35076	Adc35076 Human bre
4	113	87.6	1937	4 ABB58985	Abb58985 Drosophil
5	110	85.3	140	4 AAU16619	Aau16619 Human nov
6	110	85.3	140	6 ABU56688	Abu56688 Human nov
7	110	85.3	235	4 AAU16205	Aau16205 Human nov
8	110	85.3	235	6 ABUS5275	Abu5275 Human nov
9	107	82.9	688	3 AAY57898	Aay57898 Human tra
10	107	82.9	1390	5 ABU65152	Abu65152 Human NOV
11	100	77.5	513	4 ABB68341	Abb68341 Drosophil
12	96	74.4	947	2 AAU81168	Aau81168 Transcrip
13	96	74.4	947	2 AAY07114	Aay07114 W09904265
14	93	72.1	124	3 AAB32659	Aab32659 Eucalyptu
15	92	71.3	714	4 ABB10109	Abb10109 Rice deri
16	88	68.2	293	3 AAB32946	Aab32946 Pinus rad
17	88	68.2	422	3 AAG41764	Aag41764 Arabidops
18	88	68.2	568	3 AAG41763	Aag41763 Arabidops
19	88	68.2	590	3 AAG41762	Aag41762 Arabidops
20	88	68.2	2414	2 AAR84882	Aar84882 Transcrip
21	88	68.2	2414	2 AAU40057	Aau40057 Cellular
22	88	68.2	2414	5 ABB06340	Abb06340 Human p30
23	88	68.2	2414	6 ABU03968	Abu03968 Human exp
24	88	68.2	2414	6 ABU03980	Abu03980 Human exp
25	88	68.2	2414	6 ABU03975	Abu03975 Human exp

26	88	68.2	2414	6 ABU03978	Abu03978 Human exp
27	88	68.2	2414	6 ABU03976	Abu03976 Human exp
28	88	68.2	2441	2 AAR79054	Aar79054 CREB bind
29	88	68.2	2441	2 AAU40058	Aau40058 Cellular
30	88	68.2	2441	3 AAY94252	Aay94252 Mouse nuc
31	88	68.2	2441	4 ABB44555	Abb44555 Mouse wou
32	88	68.2	2441	6 ABU03977	Abu03977 Human exp
33	88	68.2	2442	4 ABB44556	Abb44556 Human wou
34	88	68.2	2442	5 ABG95114	Abg95114 Human tra
35	88	68.2	2442	6 ABU03979	Abu03979 Human exp
36	88	68.2	2442	6 ABU03981	Abu03981 Human exp
37	88	68.2	2442	6 ABU03967	Abu03967 Human exp
38	88	68.2	2442	6 ABU03965	Abu03965 Human exp
39	86	66.7	330	6 ABR41589	Ab41589 Human DIT
40	86	66.7	494	3 AAG36673	Aag36673 Arabidops
41	86	66.7	521	3 AAG36672	Aag36672 Arabidops
42	86	66.7	573	6 ABR41326	Ab41326 Human DIT
43	86	66.7	620	3 AAG36671	Aag36671 Arabidops
44	86	66.7	3190	4 AAB84634	Aab84634 Amino aci
45	86	66.7	3275	4 ABB70437	Abb70437 Drosophil

ALIGNMENTS

RESULT 1  
AAB95802  
ID AAB95802 standard; protein; 303 AA.  
XX AAB95802;  
XX 26-JUN-2001 (first entry)  
XX Human protein sequence SEQ ID NO:18783.  
XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.  
XX Homo sapiens.  
XX EP1074617-A2.  
XX 07-FEB-2001.  
XX 28-JUL-2000; 2000EP-00116126.  
XX 29-JUL-1999; 95JP-00248036.  
XX 27-AUG-1999; 95JP-00300253.  
XX 11-JAN-2000; 2000JP-00118776.  
XX 02-MAY-2000; 2000JP-00183767.  
XX 09-JUN-2000; 2000JP-00241899.  
XX (HELI-) HELIX RES INST.  
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
XX WPI; 2001-318749/34.

Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.  
Claim 8; SEQ ID NO 18783; 2537pp + Sequence Listing; English.  
The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dr primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the nucleotide sequences comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end

CC sequence and an oligonucleotide comprising a sequence complementary to a  
 CC polynucleotide which comprises a 3'-end sequence, where the  
 CC oligonucleotide comprises at least 15 nucleotides and the combination of  
 CC the 5'-end sequence/3'-end sequence is selected from those defined in the  
 CC specification. The primer sets can be used in antisense therapy and in  
 CC gene therapy. The primers are useful for synthesising polynucleotides,  
 CC particularly full-length cDNAs. The primers are also useful for the  
 CC detection and/or diagnosis of the abnormality of the proteins encoded by  
 CC the full-length cDNAs. The primers allow obtaining of the full-length  
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893  
 CC represent human amino acid sequences; and AAH13629 to AAH13632 represent  
 CC present invention  
 XX  
 SQ Sequence 303 AA;

Query Match 100.0%; Score 129; DB 4; Length 303;  
 Best Local Similarity 100.0%; Pred. No. 5.9e-12;  
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRQPVDAVKLGDPYHKIIKQPM 24  
 DB 102 PRQPVDAVKLGDPYHKIIKQPM 125

RESULT 2  
 AAY07027  
 ID AAY07027 standard; protein; 754 AA.

AC AAY07027;  
 DT 02-JUL-1999 (first entry)  
 DE Breast cancer associated antigen precursor sequence.  
 KW Cancer associated antigen; diagnosis; research; treatment; human;  
 KW breast cancer; colon cancer; gastric cancer; renal cancer; lung cancer;  
 KW prostate cancer.

OS Homo sapiens.

PN WO9904265-A2.

PD 28-JAN-1999.

PF 15-JUL-1998; 98WO-US014679.

PR 17-JUL-1997; 97US-00896164.

PR 10-OCT-1997; 97US-0061599P.

PR 10-OCT-1997; 97US-0061765P.

PR 10-OCT-1997; 97US-00948705.

PR 11-OCT-1997; 97GB-00021697.

PR 22-JUN-1998; 98US-00102322.

XX (LUDW-) LUDWIG INST CANCER RES.

XX Old LJ, Scanlan MJ, Stockert E, Gure A, Chen Y, Gout I;

PI O'hare M, Obata Y, Pfrendschuh M, Tureci O, Sahin U;

XX WPI; 1999-132448/11.

XX New isolated cancer associated nucleic acids and polypeptides - isolated  
 PT using sera from cancer patients, used to develop products for the  
 PT diagnosis, monitoring or treatment of cancers.

PS Disclosure; Page 404-405; 787pp; English.

XX The invention relates to a method for diagnosing a disorder characterised  
 CC by expression of a human cancer associated antigen precursor coded for by  
 CC a nucleic acid molecule (NAM). The method comprises: (a) contacting a  
 CC biological sample isolated from a subject with an agent that specifically  
 CC binds to the NAM, an expression product or a fragment of an expression

CC product complexed with an HLA molecule; and (b) determining the  
 CC interaction between the agent and the NAM or the expression product as a  
 CC determination of the disorder. The products and methods can be used in  
 CC the diagnosis, monitoring, research, or treatment of conditions  
 CC characterised by the expression of various cancer associated antigens.  
 CC The invention provides nucleic acid sequences and encoded polypeptides  
 CC which are cancer associated antigen precursors expressed in human breast  
 CC cancer, renal cancer, colon cancer, gastric cancer, prostate cancer and  
 CC lung cancer

XX Sequence 754 AA;

Query Match 100.0%; Score 129; DB 2; Length 754;  
 Best Local Similarity 100.0%; Pred. No. 1.7e-11;  
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRQPVDAVKLGDPYHKIIKQPM 24  
 DB 52 PRQPVDAVKLGDPYHKIIKQPM 75

RESULT 3

ADC35076  
 ID ADC35076 standard; protein; 801 AA.

AC ADC35076;

DT 18-DEC-2003 (first entry)

DE Human breast cancer antigen seq id 42.

XX breast cancer; breast cancer diagnosis; breast cancer antigen.

OS Homo sapiens.

PN US2003108888-A1.

PD 12-JUN-2003.

PF 15-MAY-2002; 2002US-00146473.

PR 15-MAY-2001; 2001US-0291150P.

XX (LUDW-) LUDWIG INST CANCER RES.

PI Scanlan MJ, Gout I, Stockert E, Old LJ, Gure A, Chen Y;

XX WPI; 2003-829397/77.

DR N-PSDB; ADC35118.

XX Diagnosing breast cancer in subject by obtaining biological sample from  
 PT subject, contacting sample with breast cancer-associated polypeptides,  
 PT determining specific binding between polypeptides and agents in sample.

PS Example 2; SEQ ID NO 42; 173pp; English.

XX The invention describes a method of diagnosing breast cancer in subject  
 CC comprising contacting biological sample from subject with at least two  
 CC different breast cancer-associated polypeptides (I) encoded by nucleic  
 CC acid molecules (II) comprising sequence chosen from 42 fully defined  
 CC sequences as given in specification, determining specific binding between  
 CC (I) and agents in sample, where presence of the binding is diagnostic for  
 CC breast cancer. The method is useful for diagnosing breast cancer in a  
 CC subject. The sample is blood, lymph node fluid or breast discharge fluid.  
 CC This is the amino acid sequence of a breast cancer antigen.

XX Sequence 801 AA;

Query Match 100.0%; Score 129; DB 7; Length 801;  
 Best Local Similarity 100.0%; Pred. No. 1.8e-11;  
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRQPVDAVKLGDPYHKIIKQPM 24

Db 99 FRQPDVAVKGLPDYHKIIKQPM 122

RESULT 4  
ABB58985  
ID ABB58985 standard; protein; 1937 AA.

XX AC ABB58985;  
XX DT 26-MAR-2002 (first entry)

XX DE Drosophila melanogaster polypeptide SEQ ID NO 3747.

XX KW Drosophila; developmental biology; cell signalling; insecticide;  
XX KW pharmaceutical.

XX OS Drosophila melanogaster.

XX PN WO200171042-A2.

XX PD 27-SEP-2001.

XX PF 23-MAR-2001; 2001WO-US009231.

XX PR 23-MAR-2000; 2000US-0191637P.

XX PR 11-JUL-2000; 2000US-00614150.

XX PA (PEKE ) PE CORP NY.

XX PI Venter JC, Adams M, Li PWD, Myers EW;

XX DR WPI; 2001-656860/75.

XX DR N-PSDB; ABL03089.

XX PT New isolated nucleic acid detection reagent for detecting 1000 or more  
XX PT genes from drosophila and for elucidating cell signalling and cell-cell  
XX PT interactions.

XX PS Disclosure; SEQ ID NO 3747; 21pp + Sequence Listing; English.

XX CC The invention relates to an isolated nucleic acid detection reagent  
XX CC capable of detecting 1000 or more genes from Drosophila. The invention is  
XX CC useful in developmental biology and in elucidating cell signalling and  
XX CC cell-cell interactions in higher eukaryotes for the development of  
XX CC insecticides, therapeutics and pharmaceutical drugs. The invention  
XX CC discloses genomic DNA sequences (ABU16176-ABU30511), expressed DNA  
XX CC sequences (ABU1840-ABU16175) and the encoded proteins (ABB57737-  
XX CC ABB72072). The sequence data for this patent did not form part of the  
XX CC printed specification, but was obtained in electronic format directly  
XX CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 1937 AA;

Query Match 87.6%; Score 113; DB 4; Length 1937;  
Best Local Similarity 87.5%; Pred. No. 1.8e-09;  
Matches 21; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FRQPDVAVKGLPDYHKIIKQPM 24  
Db 59 FQPDVAVKGLPDYHKIIKQPM 82

RESULT 5  
AAU16619  
ID AAU16619 standard; protein; 140 AA.

XX AC AAU16619;

XX DT 07-NOV-2001 (first entry)

XX DE Human novel secreted protein, Seq ID 1572.

KW Human; immunosuppressive; antiarthritic; antirheumatic; cytostatic;  
KW cardiant; vasotropic; cerebroprotective; nootropic; neuroprotective;  
KW antibacterial; virucide; fungicide; ophthalmological; vulnerary;  
KW secreted protein; rheumatoid arthritis; hyperproliferative disorder;  
KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;  
KW cerebral ischaemia; angiogenesis; nervous system disorder;  
KW Alzheimer's disease; infection; ocular disorder; corneal infection;  
KW wound healing; epithelial cell proliferation; skin ageing; food additive;  
KW preservative; antiproliferative.

XX OS Homo sapiens.

XX PN WO200155322-A2.

XX PD 02-AUG-2001.

XX PF 17-JAN-2001; 2001WO-US001341.

XX PR 31-JAN-2000; 2000US-0179065P.

XX PR 04-FEB-2000; 2000US-0180628P.

XX PR 24-FEB-2000; 2000US-0184684P.

XX PR 02-MAR-2000; 2000US-0186350P.

XX PR 16-MAR-2000; 2000US-0189874P.

XX PR 17-MAR-2000; 2000US-0190076P.

XX PR 18-APR-2000; 2000US-0198123P.

XX PR 19-MAY-2000; 2000US-0205515P.

XX PR 07-JUN-2000; 2000US-0209457P.

XX PR 28-JUN-2000; 2000US-0214886P.

XX PR 30-JUN-2000; 2000US-0215135P.

XX PR 07-JUL-2000; 2000US-0216647P.

XX PR 11-JUL-2000; 2000US-0216880P.

XX PR 07-JUL-2000; 2000US-0217487P.

XX PR 11-JUL-2000; 2000US-0217496P.

XX PR 14-JUL-2000; 2000US-0218290P.

XX PR 26-JUL-2000; 2000US-0220963P.

XX PR 26-JUL-2000; 2000US-0220964P.

XX PR 14-AUG-2000; 2000US-0224518P.

XX PR 14-AUG-2000; 2000US-0224519P.

XX PR 14-AUG-2000; 2000US-0225213P.

XX PR 14-AUG-2000; 2000US-0225214P.

XX PR 14-AUG-2000; 2000US-0225214P.

XX PR 14-AUG-2000; 2000US-0225267P.

XX PR 14-AUG-2000; 2000US-0225268P.

XX PR 14-AUG-2000; 2000US-0225270P.

XX PR 14-AUG-2000; 2000US-0225447P.

XX PR 14-AUG-2000; 2000US-0225757P.

XX PR 14-AUG-2000; 2000US-0225758P.

XX PR 14-AUG-2000; 2000US-0225759P.



XX OS Homo sapiens.  
 XX PN US2002132753-A1.  
 XX PD 19-SEP-2002.  
 XX PF 17-JAN-2001; 2001US-00764864.  
 XX PR 31-JAN-2000; 2000US-0179065P.  
 XX PR 04-FEB-2000; 2000US-0180628P.  
 XX PR 28-JUN-2000; 2000US-0214886P.  
 XX PR 07-JUL-2000; 2000US-0216647P.  
 XX PR 07-JUL-2000; 2000US-0216880P.  
 XX PR 11-JUL-2000; 2000US-0217487P.  
 XX PR 11-JUL-2000; 2000US-0217496P.  
 XX PR 14-JUL-2000; 2000US-0218290P.  
 XX PR 26-JUL-2000; 2000US-0220963P.  
 XX PR 26-JUL-2000; 2000US-0220964P.  
 XX PR 14-AUG-2000; 2000US-0224518P.  
 XX PR 14-AUG-2000; 2000US-0224519P.  
 XX PR 14-AUG-2000; 2000US-0225267P.  
 XX PR 14-AUG-2000; 2000US-0225368P.  
 XX PR 14-AUG-2000; 2000US-0225370P.  
 XX PR 14-AUG-2000; 2000US-0225447P.  
 XX PR 14-AUG-2000; 2000US-0225757P.  
 XX PR 14-AUG-2000; 2000US-0225758P.  
 XX PR 22-AUG-2000; 2000US-0226868P.  
 XX PR 30-AUG-2000; 2000US-0228924P.  
 XX PR 01-SEP-2000; 2000US-0229287P.  
 XX PR 01-SEP-2000; 2000US-0229343P.  
 XX PR 01-SEP-2000; 2000US-0229344P.  
 XX PR 01-SEP-2000; 2000US-0229345P.  
 XX PR 05-SEP-2000; 2000US-0229309P.  
 XX PR 05-SEP-2000; 2000US-0229513P.  
 XX PR 08-SEP-2000; 2000US-0231413P.  
 XX PR 21-SEP-2000; 2000US-0234823P.  
 XX PR 21-SEP-2000; 2000US-0234274P.  
 XX PR 21-SEP-2000; 2000US-0234977P.  
 XX PR 27-SEP-2000; 2000US-0235834P.  
 XX PR 27-SEP-2000; 2000US-0236327P.  
 XX PR 29-SEP-2000; 2000US-0236367P.  
 XX PR 29-SEP-2000; 2000US-0236368P.  
 XX PR 29-SEP-2000; 2000US-0236369P.  
 XX PR 29-SEP-2000; 2000US-0236370P.  
 XX PR 29-SEP-2000; 2000US-0236802P.  
 XX PR 02-OCT-2000; 2000US-0237037P.  
 XX PR 02-OCT-2000; 2000US-0237038P.  
 XX PR 02-OCT-2000; 2000US-0237039P.  
 XX PR 02-OCT-2000; 2000US-0237040P.  
 XX PR 13-OCT-2000; 2000US-0239935P.  
 XX PR 20-OCT-2000; 2000US-0240960P.  
 XX PR 20-OCT-2000; 2000US-0241785P.  
 XX PR 20-OCT-2000; 2000US-0241809P.  
 XX PR 01-NOV-2000; 2000US-0244617P.  
 XX PR 17-NOV-2000; 2000US-0249299P.  
 XX PR 08-DEC-2000; 2000US-0251856P.  
 XX PR 08-DEC-2000; 2000US-0251868P.  
 XX PR 08-DEC-2000; 2000US-0251869P.  
 XX PR (ROSE/) ROSEN C A.  
 XX PR (RUBE/) RUBEN S M.  
 XX PR (BARA/) BARASH S C.  
 XX PI Rosen CA, Ruben SM, Barash SC;  
 XX DR WPI; 2003-147444/14.  
 XX DR N-ESDB; ABX73947.  
 XX PT New polypeptides and nucleic acids, useful in gene therapy for treating,  
 XX PT inhibiting or preventing e.g. neural, immune system, muscular,  
 XX PT respiratory, reproductive, gastrointestinal, pulmonary, cardiovascular or  
 XX PT renal disorders.

XX PS Claim 11; SEQ ID NO 1572; 402pp; English.  
 XX CC The invention relates to human novel polypeptides and their associated  
 CC polynucleotides. The polypeptides and polynucleotides are useful in gene  
 CC therapy for treating, inhibiting or preventing neural disorders, immune  
 CC system disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis  
 CC and multiple sclerosis), muscular disorders, respiratory diseases (e.g.  
 CC nasal vestibulitis, nasal polyps and sinusitis), reproductive disorders,  
 CC gastrointestinal disorders, pulmonary disorders, cardiovascular disorders  
 CC (e.g. congenital heart defects, Ebstein's anomaly and hypoplastic left  
 CC heart syndrome), renal disorders (e.g. acute kidney failure and end-stage  
 CC renal disease), hyperproliferative disorders (e.g. Hodgkin's disease and  
 CC leukaemia), inflammatory diseases (e.g. septic shock, Bursitis and  
 CC appendicitis), allergic reactions and conditions (e.g. asthma), blood  
 CC related disorders (e.g. thrombosis, atherosclerosis and myocardial  
 CC infarction) and cancerous diseases. Sequences ABU54914-ABU55699 and  
 CC ABU55748 represent human novel polypeptides of the invention  
 XX SQ Sequence 140 AA;  
 SQ Query Match 85.3%; Score 110; DB 6; Length 140;  
 Best Local Similarity 83.3%; Pred. No. 2.5e-09;  
 Matches 20; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 FRQPDVAVKLGIPDYHKIKQPM 24  
 DB 33 FYQPDVAVKLGIPDYHKIKQPM 56  
 RESULT 7  
 AAU16206  
 ID AAU16206 standard; protein; 235 AA.  
 XX AC AAU16206;  
 XX DT 07-NOV-2001 (first entry)  
 XX DE Human novel secreted protein, Seq ID 1159.  
 XX KW Human; immunosuppressive; antiarthritic; antirheumatic; cytostatic;  
 KW cardiant; vasotropic; cerebroprotective; neurotropic; neuroprotective;  
 KW antibacterial; virucide; fungicide; ophthalmological; vulnerary;  
 KW secreted protein; rheumatoid arthritis; hyperproliferative disorder;  
 KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;  
 KW cerebral ischaemia; angiogenesis; nervous system disorder;  
 KW Alzheimer's disease; infection; ocular disorder; corneal infection;  
 KW wound healing; epithelial cell proliferation; skin ageing; food additive;  
 KW preservative; antiproliferative.  
 XX OS Homo sapiens.  
 XX PN WO200155322-A2.  
 XX PD 02-AUG-2001.  
 XX PF 17-JAN-2001; 2001WO-US001341.  
 XX PR 31-JAN-2000; 2000US-0179065P.  
 XX PR 04-FEB-2000; 2000US-0180628P.  
 XX PR 24-FEB-2000; 2000US-0184664P.  
 XX PR 02-MAR-2000; 2000US-0186350P.  
 XX PR 16-MAR-2000; 2000US-0189874P.  
 XX PR 17-MAR-2000; 2000US-0190076P.  
 XX PR 18-APR-2000; 2000US-0198123P.  
 XX PR 19-MAY-2000; 2000US-0205515P.  
 XX PR 07-JUN-2000; 2000US-0209467P.  
 XX PR 28-JUN-2000; 2000US-0214886P.  
 XX PR 30-JUN-2000; 2000US-0215135P.  
 XX PR 07-JUL-2000; 2000US-0216647P.  
 XX PR 11-JUL-2000; 2000US-0216880P.  
 XX PR 11-JUL-2000; 2000US-0217487P.  
 XX PR 11-JUL-2000; 2000US-0217496P.





XX AAY57898;  
AC 23-MAR-2000 (first entry)  
DE Human transmembrane protein HTMPN-22.  
XX  
XX Human; transmembrane protein; HTMPN; diagnosis; immunospecific;  
KW antiproliferative; neuroprotective; immune disorder;  
XX reproductive disorder; smooth muscle disorder; neurological disorder;  
KW gastrointestinal disorder; developmental disorder;  
KW cell proliferative disorder.  
XX  
XX Homo sapiens.  
OS  
XX WO9961471-A2.  
XX  
XX 02-DEC-1999.  
XX  
XX 28-MAY-1999; 99WO-US011904.  
XX  
XX 29-MAY-1998; 98US-0087260P.  
XX 02-JUL-1998; 98US-0091674P.  
XX 02-OCT-1998; 98US-0102854P.  
XX 24-NOV-1998; 98US-0109863P.  
XX  
XX (INCY-) INCYTE PHARM INC.  
XX  
XX Tang YT, Lal P, Hillman JL, Yue H, Guegler KJ, Corley NC;  
PI Bandman O, Patterson C, Gorgone GA, Kaser MR, Baughn MR, Au-Young J;  
XX  
XX WPI; 2000-072605/06.  
DR N-PSDB; AAZ56719.  
XX  
XX Proteins, polynucleotides, vectors, host cells and antibodies used to  
PT diagnose, treat or prevent immune, reproductive, smooth muscle,  
PT neurological, gastrointestinal, developmental and cell proliferative  
PT disorders.  
XX  
XX Claim 1; Page 128-130; 229pp; English.  
XX  
XX AAZ56698 to AAZ56776 encode AAY57877 to AAY57955 which represent human  
CC transmembrane proteins designated HTMPN-1 to HTMPN-79, respectively. The  
CC transmembrane protein have immunospecific, antiproliferative and  
CC neuroprotective activities. The human transmembrane proteins,  
CC polynucleotides encoding them and other compositions and methods from the  
CC present invention, can be used for the diagnosis, treatment or prevention  
CC of immune, reproductive, smooth muscle, neurological, gastrointestinal,  
CC developmental and cell proliferative disorders. The HTMPN's can be used  
CC to treat or prevent disorders associated with a decreased expression or  
CC activity of HTMPN  
XX  
XX Sequence 688 AA;  
SQ  
Query Match 82.9%; Score 107; DB 3; Length 688;  
Best Local Similarity 83.3%; Pred. No. 4.7e-08;  
Matches 20; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 PROPVDKVLGPDYKIIKQPM 24  
Db 83 FQPPVDKVLGPDYKIIKTPMD 106  
RESULT 10  
ABU65152  
ID ABU65152 standard; protein; 1390 AA.  
XX  
XX ABU65152;  
AC  
XX 20-MAY-2003 (first entry)  
DT  
XX Human NOV79a protein.  
DE  
XX

KW NOVX; cytostatic; cardiant; antiarteriosclerotic; antiasthmatic; cancer;  
KW hypotensive; cardiomyopathy; bronchial asthma; gene therapy; vaccine;  
human.  
XX  
XX Homo sapiens.  
XX  
XX WO200272757-A2.  
XX  
XX 19-SEP-2002.  
XX  
XX 08-MAR-2002; 2002WO-US006908.  
XX  
XX 08-MAR-2001; 2001US-0274101P.  
XX 08-MAR-2001; 2001US-0274194P.  
XX 08-MAR-2001; 2001US-0274251P.  
XX 08-MAR-2001; 2001US-0274322P.  
XX 09-MAR-2001; 2001US-0274849P.  
XX 12-MAR-2001; 2001US-0275235P.  
XX 13-MAR-2001; 2001US-0275578P.  
XX 13-MAR-2001; 2001US-0275579P.  
XX 13-MAR-2001; 2001US-0275601P.  
XX 14-MAR-2001; 2001US-0276000P.  
XX 16-MAR-2001; 2001US-0276776P.  
XX 19-MAR-2001; 2001US-0276994P.  
XX 20-MAR-2001; 2001US-0277239P.  
XX 20-MAR-2001; 2001US-0277321P.  
XX 20-MAR-2001; 2001US-0277327P.  
XX 21-MAR-2001; 2001US-0277919P.  
XX 22-MAR-2001; 2001US-0277833P.  
XX 23-MAR-2001; 2001US-0278152P.  
XX 26-MAR-2001; 2001US-0278894P.  
XX 27-MAR-2001; 2001US-0278999P.  
XX 27-MAR-2001; 2001US-0279036P.  
XX 28-MAR-2001; 2001US-0279344P.  
XX 30-MAR-2001; 2001US-0277388P.  
XX 30-MAR-2001; 2001US-0279959P.  
XX 30-MAR-2001; 2001US-0280233P.  
XX 02-APR-2001; 2001US-0280802P.  
XX 02-APR-2001; 2001US-0280822P.  
XX 02-APR-2001; 2001US-0280900P.  
XX 04-APR-2001; 2001US-0281194P.  
XX 13-APR-2001; 2001US-0283675P.  
XX 30-APR-2001; 2001US-0287424P.  
XX 02-MAY-2001; 2001US-0288066P.  
XX 03-MAY-2001; 2001US-0288342P.  
XX 03-MAY-2001; 2001US-0288528P.  
XX 15-MAY-2001; 2001US-0291190P.  
XX 16-MAY-2001; 2001US-0291099P.  
XX 16-MAY-2001; 2001US-0291240P.  
XX 30-MAY-2001; 2001US-0294485P.  
XX 31-MAY-2001; 2001US-0294899P.  
XX 31-MAY-2001; 2001US-0294899P.  
XX 18-JUN-2001; 2001US-0299027P.  
XX 19-JUN-2001; 2001US-0299303P.  
XX 19-JUN-2001; 2001US-0299310P.  
XX 10-JUL-2001; 2001US-0304354P.  
XX 31-JUL-2001; 2001US-0309198P.  
XX 16-AUG-2001; 2001US-0312903P.  
XX 10-SEP-2001; 2001US-0318462P.  
XX 12-SEP-2001; 2001US-0318770P.  
XX 27-SEP-2001; 2001US-0325430P.  
XX 27-SEP-2001; 2001US-0325681P.  
XX 18-OCT-2001; 2001US-0330380P.  
XX 31-OCT-2001; 2001US-0335301P.  
XX 14-NOV-2001; 2001US-0332172P.  
XX 14-NOV-2001; 2001US-0332271P.  
XX 14-NOV-2001; 2001US-0332272P.  
XX 14-NOV-2001; 2001US-0333184P.  
XX 14-NOV-2001; 2001US-0333272P.  
XX 21-NOV-2001; 2001US-0332094P.  
XX 03-DEC-2001; 2001US-0337426P.  
XX 03-DEC-2001; 2001US-0338092P.  
XX 04-DEC-2001; 2001US-0337185P.



```

PR 03-JAN-2002; 2002US-0345705P.
PR 07-MAR-2002; 2002US-00092900.
XX (CURA-) CURAGEN CORP.
XX
PI Padigar M, Spytek KA, Shenoy SG, Taupier RJ, Pena CEA, Li L;
PI Zehrhusen BD, Gusev V, Ji W, Gorman L, Miller CE, Kekuda R;
PI Pattarajan M, Gangolli E, Vernet CAM, Guo X, Tchernev V;
PI Fernandes ER, Casman SJ, Malyankar UM, Gerlach V, Liu Y, Anderson D;
PI Spaderna SK, Catterton E, Burgess C, Leite M, Zhong H, Alsobrook JP;
PI Lepley DM, Rieger DK;
XX WPI; 2002-723332/78.
DR N-PSDB; ABX97119.
XX
XX NOVX polypeptides and polynucleotides, useful for preventing or treating
PT a disorder associated with aberrant NOVX expression or activity e.g.,
PT cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial
PT asthma.
XX
XX Claim 1; Page 323; 1103pp; English.
XX
XX This invention describes novel human NOVX polypeptides which have
CC cytostatic, cardiac, antiarteriosclerotic, antiasthmatic and hypotensive
CC activity. Pharmaceutical compositions comprising the NOVX proteins or
CC nucleic acid molecules or NOVX antibodies are useful for preventing or
CC treating a disorder associated with aberrant NOVX expression or activity
CC e.g. cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial
CC asthma. The products of the invention can be used for gene therapy or in
CC a vaccine. ABU65041-ABU65218 represent the NOVX polypeptides encoded by
CC ABX97008-ABX97185
XX
XX Sequence 1390 AA;
SQ
Query Match 82.9%; Score 107; DB 5; Length 1390;
Best Local Similarity 83.3%; Pred. No. 1.1e-07;
Matches 20; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 FROPVDAVKLGDPYHKIIKPMD 24
DB 83 FQFPVDAVKLNLDPYKIIKTPMD 106
RESULT 11
ABX68341
ID ABB68341 standard; protein; 513 AA.
XX
XX ABB68341;
AC
XX 26-MAR-2002 (first entry)
DT
XX Drosophila melanogaster polypeptide SEQ ID NO 31815.
DE
XX Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical.
XX
XX Drosophila melanogaster.
OS
XX WO200171042-A2.
PN
XX 27-SEP-2001.
PD
XX
XX 23-MAR-2001; 2001WO-US0009231.
PF
XX 23-MAR-2000; 2000US-0191637P.
PR
XX 11-JUL-2000; 2000US-00614150.
XX
XX (PEKE ) PE CORP NY.
PA
XX Venter JC, Adams M, Li PWD, Myers EW;
PI
XX WPI; 2001-656860/75.
DR
XX N-PSDB; ABL12444.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions.
XX
XX Disclosure; SEQ ID NO 31815; 21pp + Sequence Listing; English.
PS
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 513 AA;
SQ
Query Match 77.5%; Score 100; DB 4; Length 513;
Best Local Similarity 66.7%; Pred. No. 4.4e-07;
Matches 16; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
QY 1 FROPVDAVKLGDPYHKIIKPMD 24
DB 63 FRHPVDSVSLGVPDYHAVVRHPMD 86
RESULT 12
AAW81168
ID AAW81168 standard; protein; 947 AA.
XX
XX AAW81168;
AC
XX 05-MAR-1999 (first entry)
DT
XX Transcriptional regulatory factor RING3.
DE
XX Human; transcriptional regulatory factor; RING3; TSB; cancer;
XX testis specific bromodomain; testicular cell proliferation.
XX
XX Homo sapiens.
OS
XX WO9848015-A1.
PN
XX 29-OCT-1998.
PD
XX
XX 17-APR-1998; 98WO-JP001782.
PF
XX 18-APR-1997; 97JP-00116402.
PR
XX (CHUS ) CHUGAI RES INST MOLECULAR MEDICINE INC.
PA
XX Jones MH;
PI
XX WPI; 1998-583658/49.
DR
XX N-PSDB; AAV68343.
XX
XX Transcriptional regulator gene containing bromodomain sequence - may be
PT expressed in testis tissue and is useful in treatment of cancer and other
PT proliferative disorders.
XX
XX Claim 1; Page 19-24; 42pp; Japanese.
PS
XX The present sequence represents the human transcriptional regulatory
CC factor RING3, which is isolated from testicular cells. RING3 contains a
CC testis specific bromodomain (TSB) which is expressed specifically in
CC testis tissue and also expressed in certain tumour lines. The transgenic
CC cells may be used to express RING3 which is a TSB expression protein. The
CC TSB expression product can be used in the treatment of cancer and other
CC proliferative disorders, and in screening of compounds for ability to
CC bind to it (e.g. for use as drugs by modulation of transcriptional

```

CC regulation). DNA capable of hybridising to RING3 polynucleotides may be  
 CC used for construction of probes and primers

XX SQ Sequence 947 AA;  
 Query Match 74.4%; Score 96; DB 2; Length 947;  
 Best Local Similarity 75.0%; Pred. No. 3.8e-06;  
 Matches 18; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 FRQPVDAVKLGPDYHKKIKQPM 24  
 DB 52 FQRPVDAVKLGPDYHKKIKQPM 75

RESULT 13  
 AAY07114  
 ID AAY07114 standard; protein; 947 AA.

XX AC AAY07114;  
 XX DT 02-JUL-1999 (first entry)  
 XX DE WO9904265 Seq ID No: 685.  
 XX KW Cancer associated antigen; diagnosis; research; treatment; human;  
 XX KW breast cancer; colon cancer; gastric cancer; renal cancer; lung cancer;  
 XX KW prostate cancer.

XX OS Homo sapiens.  
 XX PN WO9904265-A2.

XX PD 28-JAN-1999.  
 XX PF 15-JUL-1998; 98WO-US014679.

XX PR 17-JUL-1997; 97US-00896164.  
 XX PR 10-OCT-1997; 97US-00611599P.  
 XX PR 10-OCT-1997; 97US-0061765P.  
 XX PR 10-OCT-1997; 97US-00948705.  
 XX PR 11-OCT-1997; 97GB-00021697.  
 XX PR 22-JUN-1998; 98US-00102322.

XX PA (LUDW-) LUDWIG INST CANCER RES.

XX PI Old LJ, Scanlan MJ, Stockert E, Gure A, Chen Y, Gout I;  
 XX PI O'hare M, Obata Y, Pfrendschuh M, Tureci O, Sahin U;

XX DR WPI; 1999-132448/11.

XX PT New isolated cancer associated nucleic acids and polypeptides - isolated  
 XX PT using sera from cancer patients, used to develop products for the  
 XX PT diagnosis, monitoring or treatment of cancers.

XX PS Disclosure; Page 728-730; 787pp; English.

XX CC The invention relates to a method for diagnosing a disorder characterised  
 XX CC by expression of a human cancer associated antigen precursor coded for by  
 XX CC a nucleic acid molecule (NAM). The method comprises: (a) contacting a  
 XX CC biological sample isolated from a subject with an agent that specifically  
 XX CC binds to the NAM, an expression product or a fragment of an expression  
 XX CC product complexed with an HLA molecule; and (b) determining the  
 XX CC interaction between the agent and the NAM or the expression product as a  
 XX CC determination of the disorder. The products and methods can be used in  
 XX CC the diagnosis, monitoring, research, or treatment of conditions  
 XX CC characterised by the expression of various cancer associated antigens.  
 XX CC The invention provides nucleic acid sequences and encoded polypeptides  
 XX CC which are cancer associated antigen precursors expressed in human breast  
 XX CC cancer, renal cancer, colon cancer, gastric cancer, prostate cancer and  
 XX CC lung cancer

XX SQ Sequence 947 AA;

Query Match 74.4%; Score 96; DB 2; Length 947;  
 Best Local Similarity 75.0%; Pred. No. 3.8e-06;  
 Matches 18; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 FRQPVDAVKLGPDYHKKIKQPM 24  
 DB 52 FQRPVDAVKLGPDYHKKIKQPM 75

RESULT 14  
 AAB32659  
 ID AAB32659 standard; protein; 124 AA.

XX AC AAB32659;

XX DT 25-JAN-2001 (first entry)

XX DE Eucalyptus grandis transcription factor protein sequence #117.

XX KW Plant; transcription factor; gene expression; eucalyptus; pine; acacia;  
 XX KW poplar; sweetgum; teak; mahogany; bz1P; G-box binding factor;  
 XX KW basic helix-loop-helix zipper; homeotic; homeodomain; homeobox; MADS;  
 XX KW homeodomain zipper; LIM domain; AP2; EREBs; zinc finger domain;  
 XX KW type 2 Cys2His2; CCAAT box element; MYB.

XX OS Eucalyptus grandis.

XX PN WO200053724-A2.

XX PD 14-SEP-2000.

XX PF 09-MAR-2000; 2000WO-US006112.

XX PR 11-MAR-1999; 99US-00266513.

XX PR 18-AUG-1999; 99US-0149485P.

XX PA (GENE-) GENESIS RES & DEV CORP LTD.

XX PI (FLET-) FLETCHER CHALLENGE FORESTS LTD.

XX PI Wood M, McGrath A, Shenk MA, Glenn M;

XX DR WPI; 2000-579369/54.

XX PT New isolated polynucleotide encoding a plant transcription factor for  
 XX PT producing a plant e.g. a woody plant, preferably eucalyptus or pine,  
 XX PT having modified gene expression or modified activity of a polypeptide.

XX PS Claim 8; Page 246; 747pp; English.

XX CC The present invention relates to novel plant transcription factors from  
 XX CC Eucalyptus grandis or Pinus radiata. The present sequence is one such  
 XX CC transcription factor. The transcription factor may be used to produce a  
 XX CC plant having modified gene expression such as a woody plant e.g. a  
 XX CC eucalyptus, pine, acacia, poplar, sweetgum, teak, or mahogany species or  
 XX CC to modify the activity of a polypeptide in a plant. The transcription  
 XX CC factors of the present invention are members from the following families  
 XX CC of regulatory proteins: bz1P, bz1P family of G-box binding factors, basic  
 XX CC helix-loop-helix zipper, homeotic/homeodomain/homeobox/MADS, homeodomain  
 XX CC zipper, LIM domain, AP2 and EREBs, zinc finger domains of type 2  
 XX CC Cys2His2, CCAAT box elements and MYB

XX SQ Sequence 124 AA;

Query Match 72.1%; Score 93; DB 3; Length 124;  
 Best Local Similarity 75.0%; Pred. No. 1.1e-06;  
 Matches 18; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 FRQPVDAVKLGPDYHKKIKQPM 24  
 DB 44 FNTFVDAVGLGHDYHKKIKQPM 67

RESULT 15

ABB10109  
ID ABB10109 standard; protein; 714 AA.  
XX AC ABB10109;  
XX DT 01-JUL-2002 (first entry)  
XX DE Rice derived plant resistance protein#3.  
XX KW Plant; resistance; cerebroside type elicitor; rice blight microbe.  
XX OS Oryza sativa.  
XX PN JP2000342262-A.  
XX PD 12-DEC-2000.  
XX PF 31-MAY-1999; 99JP-00153146.  
XX PR 31-MAY-1999; 99JP-00153146.  
XX PA (SHOK-) SHOKUBUTSU SOGYO SYSTEM KENKYUSHO KK.  
XX DR WPI; 2001-268250/28.  
XX DR N-PSDB; ABL56859.  
XX PT New gene for giving acquired resistance to a plant and a method for  
PT screening for it, comprising using a cerebroside type elicitor derived  
PT from a rice blight microbe.  
XX PS Claim 3; Page 14-16; 34pp; Japanese.  
XX CC The invention relates to a gene encoding a protein giving acquired  
CC resistance to a plant, and a method for screening for it using a  
CC cerebroside type elicitor derived from a rice blight microbe. The new  
CC gene can be used to give acquired resistance to a plant. The current  
CC sequence represents a rice derived plant resistance protein  
XX SQ Sequence 714 AA;  
Query Match 71.3%; Score 92; DB 4; Length 714;  
Best Local Similarity 70.8%; Pred. No. 1.2e-05;  
Matches 17; Conservative 3; Mismatches 4; Indels 0; Gaps 0;  
QY 1 FROPDAVKLGDPDYHKIKQPMQ 24  
Db 187 FDSFVDAVKLNIPDYFQIIKKPMQ 210  
Search completed: July 27, 2004, 12:07:14  
Job time : 54 secs



GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 27, 2004, 11:58:16 ; Search time 18 Seconds  
(without alignments)  
68.835 Million cell updates/sec

Title: US-09-784-553C-19\_COPY\_29\_52  
Perfect score: 129  
Sequence: 1 FRQPDVAVKGLPDYHKIIQPM 24

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/prodata/2/iaa/5A\_COMB.pep:\*  
2: /cgn2\_6/prodata/2/iaa/5B\_COMB.pep:\*  
3: /cgn2\_6/prodata/2/iaa/5A\_COMB.pep:\*  
4: /cgn2\_6/prodata/2/iaa/5B\_COMB.pep:\*  
5: /cgn2\_6/prodata/2/iaa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/prodata/2/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	129	100.0	65	1	US-08-227-536-5
2	129	100.0	65	5	PCT-US95-04682-5
3	129	100.0	754	4	US-09-392-714-20
4	110	85.3	726	4	US-09-392-714-21
5	107	82.9	722	4	US-09-392-714-22
6	96	74.4	947	4	US-09-418-780A-1
7	96	74.4	947	4	US-09-392-714-23
8	89	69.0	61	4	US-09-418-710-42
9	88	68.2	61	4	US-09-418-710-55
10	88	68.2	2414	1	US-08-227-536-2
11	88	68.2	2414	5	PCT-US95-04682-2
12	88	68.2	2441	1	US-08-194-468-2
13	88	68.2	2441	3	US-08-961-739-2
14	88	68.2	2441	4	US-09-514-247A-8
15	88	68.2	2442	4	US-09-686-316-2
16	88	68.2	2442	4	US-09-514-247A-10
17	86	66.7	65	1	US-08-227-536-6
18	86	66.7	65	5	PCT-US95-04682-6
19	61	47.3	65	1	US-08-227-536-3
20	61	47.3	65	1	US-08-227-536-4
21	61	47.3	65	5	PCT-US95-04682-3
22	61	47.3	65	5	PCT-US95-04682-4
23	58	45.0	59	4	US-09-418-710-51
24	58	45.0	238	4	US-09-257-179-80
25	58	45.0	1872	1	US-08-188-582-14
26	58	45.0	1872	1	US-08-646-715-14
27	58	45.0	1893	1	US-08-188-582-11

28	58	45.0	1893	1	US-08-646-715-11
29	56	43.4	11	4	US-09-418-780A-14
30	56	43.4	11	4	US-09-418-780A-22
31	55	42.6	65	1	US-08-227-536-8
32	55	42.6	65	5	PCT-US95-04682-8
33	53.5	41.5	21	4	US-09-257-179-85
34	53	41.1	59	4	US-09-418-710-40
35	52	40.3	59	4	US-09-418-710-39
36	52	40.3	740	4	US-09-107-532A-7211
37	52	40.3	1969	4	US-09-418-710-72
38	52	40.3	1972	4	US-09-418-710-21
39	49	38.0	59	4	US-09-418-710-58
40	49	38.0	59	4	US-09-418-710-50
41	49	38.0	1673	4	US-09-418-710-70
42	49	38.0	1674	4	US-09-418-710-1
43	48	37.2	59	4	US-09-418-710-41
44	48	37.2	59	4	US-09-418-710-53
45	47	36.4	274	4	US-09-134-000C-5795

ALIGNMENTS

RESULT 1  
US-08-227-536-5  
; Sequence 5, Application US/08227536  
; Patent No. 5658784  
; GENERAL INFORMATION:  
; APPLICANT: Ecken, Richard  
; APPLICANT: Ewen, Mark  
; APPLICANT: Livingston, David  
; TITLE OF INVENTION: NUCLEIC ACID, ENCODING TRANSCRIPTION  
; TITLE OF INVENTION: FACTOR P300 AND USES OF P300  
; NUMBER OF SEQUENCES: 13  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Weingarten, Schurgin, Gagnebin & Hayes  
; STREET: Ten Post Office Square  
; CITY: Boston  
; STATE: MA  
; COUNTRY: US  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/227,536  
; FILING DATE: 14-APR-1994  
; CLASSIFICATION: 436  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Williams Ph.D., Kathleen A.  
; REGISTRATION NUMBER: 34,380  
; REFERENCE/DOCKET NUMBER: DPCI-308XX  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 542-2290  
; TELEFAX: (617) 451-0313  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 65 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: internal  
; US-08-227-536-5

Query Match 100.0%; Score 129; DB 1; Length 65;  
Best Local Similarity 100.0%; Pred. No. 1.3e-13;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 FRQPDVAVKGLPDYHKIIQPM 24

```
Db      6 FRQPVDAVKLGLPDYHKIIKQPM 29
|||||
RESULT 2
PCT-US95-04682-5
; Sequence 5, Application PC/TUS9504682
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: NUCLEIC ACID ENCODING TRANSCRIPTION
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Weingarten, Schurgin, Gagnebin & Hayes
; STREET: Ten Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04682
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/227,536
; FILING DATE: 14-April-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Holliday C. Heine, Ph.D.
; REGISTRATION NUMBER: 34,346
; REFERENCE/DOCKET NUMBER: DFCI-308XG999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-2290
; TELEFAX: (617) 451-0313
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 65 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHEetical: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; PCT-US95-04682-5

Query Match      100.0%; Score 129; DB 5; Length 65;
Best Local Similarity 100.0%; Pred. No. 1.3e-13;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 FRQPVDAVKLGLPDYHKIIKQPM 24
Db      6 FRQPVDAVKLGLPDYHKIIKQPM 29
|||||
RESULT 3
US-09-392-714-20
; Sequence 20, Application US/09392714A
; Patent No. 6686147
; GENERAL INFORMATION:
; APPLICANT: Scanlan, Matthew J.
; APPLICANT: Gure, Ali O.
; APPLICANT: Williamson, Barbara
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Cancer Associated Antigens and Uses
; FILE REFERENCE: L0461/7062
; CURRENT FILING DATE: 1999-09-09
; EARLIER APPLICATION NUMBER: PCT/US98/14679
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSEQ for Windows Version 3.0
; CURRENT APPLICATION NUMBER: US/09/392,714A
```

```
; CURRENT FILING DATE: 1999-09-09
; EARLIER APPLICATION NUMBER: PCT/US98/14679
; EARLIER FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 20
; LENGTH: 754
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-392-714-20
Query Match      100.0%; Score 129; DB 4; Length 754;
Best Local Similarity 100.0%; Pred. No. 2.2e-12;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 FRQPVDAVKLGLPDYHKIIKQPM 24
Db      52 FRQPVDAVKLGLPDYHKIIKQPM 75
|||||
RESULT 4
US-09-392-714-21
; Sequence 21, Application US/09392714A
; Patent No. 6686147
; GENERAL INFORMATION:
; APPLICANT: Scanlan, Matthew J.
; APPLICANT: Gure, Ali O.
; APPLICANT: Williamson, Barbara
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Cancer Associated Antigens and Uses
; FILE REFERENCE: L0461/7062
; CURRENT APPLICATION NUMBER: US/09/392,714A
; CURRENT FILING DATE: 1999-09-09
; EARLIER APPLICATION NUMBER: PCT/US98/14679
; EARLIER FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 21
; LENGTH: 726
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-392-714-21
Query Match      85.3%; Score 110; DB 4; Length 726;
Best Local Similarity 83.3%; Pred. No. 2.4e-09;
Matches 20; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy      1 FRQPVDAVKLGLPDYHKIIKQPM 24
Db      59 FYQPVDAKLNLPDYHKIIKQPM 82
|||||
RESULT 5
US-09-392-714-22
; Sequence 22, Application US/09392714A
; Patent No. 6686147
; GENERAL INFORMATION:
; APPLICANT: Scanlan, Matthew J.
; APPLICANT: Gure, Ali O.
; APPLICANT: Williamson, Barbara
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Cancer Associated Antigens and Uses
; FILE REFERENCE: L0461/7062
; CURRENT FILING DATE: 1999-09-09
; EARLIER APPLICATION NUMBER: PCT/US98/14679
; EARLIER FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSEQ for Windows Version 3.0
```

```
; SEQ ID NO 22
; LENGTH: 722
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-392-714-22

Query Match      82.9%; Score 107; DB 4; Length 722;
Best Local Similarity 83.3%; Pred. No. 7.2e-09;
Matches 20; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 FRQPVDAVKLGDPDYHKIIKQPM 24
   :|||||:|||||:|||||:|||||
Db 83 FRQPVDAVKLNLDPDYKIIKTPMD 106

RESULT 6
US-09-418-780A-1
; Sequence 1, Application US/09418780A
; Patent No. 6504009
; GENERAL INFORMATION:
; APPLICANT: Jones, Michael H.
; TITLE OF INVENTION: TRANSCRIPTIONAL REGULATOR
; FILE REFERENCE: 06501-043001
; CURRENT APPLICATION NUMBER: US/09/418,780A
; CURRENT FILING DATE: 1999-10-15
; PRIOR APPLICATION NUMBER: PCT/JP98/01782
; PRIOR FILING DATE: 1998-04-17
; PRIOR APPLICATION NUMBER: JAPAN 9/116402
; PRIOR FILING DATE: 1997-04-18
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 947
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-418-780A-1

Query Match      74.4%; Score 96; DB 4; Length 947;
Best Local Similarity 75.0%; Pred. No. 5.7e-07;
Matches 18; Conservative 3; Mismatches 3; Indels 3; Gaps 0;

QY 1 FRQPVDAVKLGDPDYHKIIKQPM 24
   :|||||:|||||:|||||:|||||
Db 52 FRQPVDAVKLGLPDYTYIIKNPMD 75

RESULT 7
US-09-392-714-23
; Sequence 23, Application US/09392714A
; Patent No. 686147
; GENERAL INFORMATION:
; APPLICANT: Scanlan, Matthew J.
; APPLICANT: Gure, Ali O.
; APPLICANT: Williamson, Barbara
; APPLICANT: Chen, Yao-Tsang
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Cancer Associated Antigens and Uses
; TITLE OF INVENTION: Therefor
; FILE REFERENCE: L0461/7062
; CURRENT APPLICATION NUMBER: US/09/392,714A
; CURRENT FILING DATE: 1999-09-09
; EARLIER APPLICATION NUMBER: PCT/US98/14679
; EARLIER FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 23
; LENGTH: 947
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-392-714-23

Query Match      74.4%; Score 96; DB 4; Length 947;
Best Local Similarity 75.0%; Pred. No. 5.7e-07;

; SEQ ID NO 22
; LENGTH: 722
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-392-714-22

Query Match      82.9%; Score 107; DB 4; Length 722;
Best Local Similarity 83.3%; Pred. No. 7.2e-09;
Matches 20; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 FRQPVDAVKLGDPDYHKIIKQPM 24
   :|||||:|||||:|||||:|||||
Db 83 FRQPVDAVKLNLDPDYKIIKTPMD 106

RESULT 6
US-09-418-780A-1
; Sequence 1, Application US/09418780A
; Patent No. 6504009
; GENERAL INFORMATION:
; APPLICANT: Jones, Michael H.
; TITLE OF INVENTION: TRANSCRIPTIONAL REGULATOR
; FILE REFERENCE: 06501-043001
; CURRENT APPLICATION NUMBER: US/09/418,780A
; CURRENT FILING DATE: 1999-10-15
; PRIOR APPLICATION NUMBER: PCT/JP98/01782
; PRIOR FILING DATE: 1998-04-17
; PRIOR APPLICATION NUMBER: JAPAN 9/116402
; PRIOR FILING DATE: 1997-04-18
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 947
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-418-780A-1

Query Match      74.4%; Score 96; DB 4; Length 947;
Best Local Similarity 75.0%; Pred. No. 5.7e-07;
Matches 18; Conservative 3; Mismatches 3; Indels 3; Gaps 0;

QY 1 FRQPVDAVKLGDPDYHKIIKQPM 24
   :|||||:|||||:|||||:|||||
Db 52 FRQPVDAVKLGLPDYTYIIKNPMD 75

RESULT 7
US-09-392-714-23
; Sequence 23, Application US/09392714A
; Patent No. 686147
; GENERAL INFORMATION:
; APPLICANT: Scanlan, Matthew J.
; APPLICANT: Gure, Ali O.
; APPLICANT: Williamson, Barbara
; APPLICANT: Chen, Yao-Tsang
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Cancer Associated Antigens and Uses
; TITLE OF INVENTION: Therefor
; FILE REFERENCE: L0461/7062
; CURRENT APPLICATION NUMBER: US/09/392,714A
; CURRENT FILING DATE: 1999-09-09
; EARLIER APPLICATION NUMBER: PCT/US98/14679
; EARLIER FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 23
; LENGTH: 947
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-392-714-23

Query Match      74.4%; Score 96; DB 4; Length 947;
Best Local Similarity 75.0%; Pred. No. 5.7e-07;

; SEQ ID NO 22
; LENGTH: 722
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-392-714-22

Query Match      82.9%; Score 107; DB 4; Length 722;
Best Local Similarity 83.3%; Pred. No. 7.2e-09;
Matches 20; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 FRQPVDAVKLGDPDYHKIIKQPM 24
   :|||||:|||||:|||||:|||||
Db 83 FRQPVDAVKLNLDPDYKIIKTPMD 106

RESULT 6
US-09-418-710-42
; Sequence 42, Application US/09418710
; Patent No. 6596482
; GENERAL INFORMATION:
; APPLICANT: Jones, Michael H.
; TITLE OF INVENTION: TRANSCRIPTIONAL REGULATOR
; FILE REFERENCE: 06501-042001
; CURRENT APPLICATION NUMBER: US/09/418,710
; CURRENT FILING DATE: 1999-10-15
; PRIOR APPLICATION NUMBER: PCT/JP98/01783
; PRIOR FILING DATE: 1998-04-17
; PRIOR APPLICATION NUMBER: JP 9/310027
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: JP 9/116570
; PRIOR FILING DATE: 1997-04-18
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 42
; LENGTH: 61
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-418-710-42

Query Match      69.0%; Score 89; DB 4; Length 61;
Best Local Similarity 66.7%; Pred. No. 3.3e-07;
Matches 16; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 FRQPVDAVKLGDPDYHKIIKQPM 24
   :|||||:|||||:|||||:|||||
Db 4 FRQPVDPQLIGIPDYEDIVKNPMD 27

RESULT 9
US-09-418-710-55
; Sequence 55, Application US/09418710
; Patent No. 6596482
; GENERAL INFORMATION:
; APPLICANT: Jones, Michael H.
; TITLE OF INVENTION: TRANSCRIPTIONAL REGULATOR
; FILE REFERENCE: 06501-042001
; CURRENT APPLICATION NUMBER: US/09/418,710
; CURRENT FILING DATE: 1999-10-15
; PRIOR APPLICATION NUMBER: PCT/JP98/01783
; PRIOR FILING DATE: 1998-04-17
; PRIOR APPLICATION NUMBER: JP 9/310027
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: JP 9/116570
; PRIOR FILING DATE: 1997-04-18
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 55
; LENGTH: 61
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-418-710-55

Query Match      68.2%; Score 88; DB 4; Length 61;
Best Local Similarity 66.7%; Pred. No. 4.7e-07;
Matches 16; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 FRQPVDAVKLGDPDYHKIIKQPM 24
   :|||||:|||||:|||||:|||||
Db 4 FRQPVDPQLIGIPDYFDIVKNPMD 27

RESULT 9
US-09-418-710-55
; Sequence 55, Application US/09418710
; Patent No. 6596482
; GENERAL INFORMATION:
; APPLICANT: Jones, Michael H.
; TITLE OF INVENTION: TRANSCRIPTIONAL REGULATOR
; FILE REFERENCE: 06501-042001
; CURRENT APPLICATION NUMBER: US/09/418,710
; CURRENT FILING DATE: 1999-10-15
; PRIOR APPLICATION NUMBER: PCT/JP98/01783
; PRIOR FILING DATE: 1998-04-17
; PRIOR APPLICATION NUMBER: JP 9/310027
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: JP 9/116570
; PRIOR FILING DATE: 1997-04-18
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 55
; LENGTH: 61
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-418-710-55

Query Match      68.2%; Score 88; DB 4; Length 61;
Best Local Similarity 66.7%; Pred. No. 4.7e-07;
Matches 16; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 FRQPVDAVKLGDPDYHKIIKQPM 24
   :|||||:|||||:|||||:|||||
Db 4 FRQPVDPQLIGIPDYFDIVKNPMD 27
```

```
RESULT 10
US-08-227-536-2
; Sequence 2, Application US/08227536
; Patent No. 5658784
; GENERAL INFORMATION:
; APPLICANT: Eckner, Richard
; APPLICANT: Ewen, Mark
; APPLICANT: Livingston, David
; TITLE OF INVENTION: NUCLEIC ACID, ENCODING TRANSCRIPTION
; TITLE OF INVENTION: FACTOR P300 AND USES OF P300
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Weingarten, Schurgin, Gagnebin & Hayes
; STREET: Ten Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/227,536
; FILING DATE: 14-APR-1994
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Ph.D., Kathleen A.
; REGISTRATION NUMBER: 34,380
; REFERENCE/DOCKET NUMBER: DFCI-308XX
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-2290
; TELEFAX: (617) 451-0313
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2414 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-227-536-2

Query Match 68.2%; Score 88; DB 1; Length 2414;
Best Local Similarity 66.7%; Pred. No. 3.2e-05;
Matches 16; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 1 FRQPDVAVKGLPDYHKIKQPM 24
Db 1075 FRQPDVAVKGLPDYHKIKQPM 1098

RESULT 12
US-08-194-468-2
; Sequence 2, Application US/08194468
; Patent No. 5750336
; GENERAL INFORMATION:
; APPLICANT: Montminy, Marc R.
; TITLE OF INVENTION: ASSAYS FOR THE IDENTIFICATION OF
; TITLE OF INVENTION: COMPOUNDS WHICH INHIBIT ACTIVATION OF CAMP AND MITOGEN
; TITLE OF INVENTION: RESPONSIVE GENES
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
; STREET: 444 South Flower Street, Suite 2000
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/194,468
; FILING DATE: 10-FEB-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Reiter, Stephen B.
; REGISTRATION NUMBER: 31,192
; REFERENCE/DOCKET NUMBER: P41 9672
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619)-546-4737
; TELEFAX: (619)-546-9392
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2441 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-194-468-2

Query Match 65.2%; Score 88; DB 1; Length 2441;
Best Local Similarity 66.7%; Pred. No. 3.3e-05;

Qy 1 FRQPDVAVKGLPDYHKIKQPM 24
Db 1075 FRQPDVAVKGLPDYHKIKQPM 1098

RESULT 11
PCT-US95-04682-2
; Sequence 2, Application PC/TUS9504682
; Patent No. 5658784
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: NUCLEIC ACID ENCODING TRANSCRIPTION
; TITLE OF INVENTION: FACTOR P300 AND USES OF P300
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Weingarten, Schurgin, Gagnebin & Hayes
; STREET: Ten Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04682
```



Matches 16; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 FRQPDVAVKLGPDYHKKIKQPM 24  
||||| :||| :|||  
Db 1112 FRQPDVQLLGIPDYFDIVKNPMD 1135

RESULT 13  
US-08-961-739-2  
; Sequence 2, Application US/08961739A  
; Patent No. 6063583  
; GENERAL INFORMATION:  
; APPLICANT: Montminy, Marc R.  
; TITLE OF INVENTION: Methods for Treating Diabetes Mellitus  
; FILE REFERENCE: SALK1650-1  
; CURRENT APPLICATION NUMBER: US/08/961,739A  
; CURRENT FILING DATE: 1997-10-31  
; EARLIER APPLICATION NUMBER: US 194,468  
; EARLIER FILING DATE: 1994-02-10  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 2441  
; TYPE: PRT  
; ORGANISM: Mus  
; FEATURE:  
; NAME/KEY: VARIANT  
; LOCATION: (1)...(2441)  
; OTHER INFORMATION: Xaa = Any Amino Acid  
US-08-961-739-2

Query Match 68.2%; Score 88; DB 3; Length 2441;  
Best Local Similarity 66.7%; Pred. No. 3.3e-05;  
Matches 16; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 FRQPDVAVKLGPDYHKKIKQPM 24  
||||| :||| :|||  
Db 1112 FRQPDVQLLGIPDYFDIVKNPMD 1135

RESULT 14  
US-09-514-247A-8  
; Sequence 8, Application US/09514247A  
; Patent No. 6365361  
; GENERAL INFORMATION:  
; APPLICANT: TANABE SEIYAKU CO. LTD.  
; APPLICANT: MIZUKAMI, Junko  
; TITLE OF INVENTION: METHOD FOR IDENTIFYING OR SCREENING AGONIST AND ANTAGONIST TO PPA  
; FILE REFERENCE: TANIGUCHI=6  
; CURRENT APPLICATION NUMBER: US/09/514,247A  
; CURRENT FILING DATE: 2000-02-28  
; PRIOR APPLICATION NUMBER: PCT/JP98/03734  
; PRIOR FILING DATE: 1998-08-24  
; PRIOR APPLICATION NUMBER: JP231084/1997  
; PRIOR FILING DATE: 1997-08-27  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 8  
; LENGTH: 2441  
; TYPE: PRT  
; ORGANISM: mouse  
US-09-514-247A-8

Query Match 68.2%; Score 88; DB 4; Length 2441;  
Best Local Similarity 66.7%; Pred. No. 3.3e-05;  
Matches 16; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 FRQPDVAVKLGPDYHKKIKQPM 24  
||||| :||| :|||  
Db 1112 FRQPDVQLLGIPDYFDIVKNPMD 1135

RESULT 15  
US-09-686-316-2  
; Sequence 2, Application US/09686316  
; Patent No. 6646115  
; GENERAL INFORMATION:  
; APPLICANT: Montminy, Marc R.  
; TITLE OF INVENTION: Methods for Treating Diabetes Mellitus  
; FILE REFERENCE: SALK1650-1  
; CURRENT APPLICATION NUMBER: US/09/686,316  
; CURRENT FILING DATE: 2000-10-10  
; PRIOR APPLICATION NUMBER: US/08/961,739  
; PRIOR FILING DATE: 1997-10-31  
; PRIOR APPLICATION NUMBER: US 194,468  
; PRIOR FILING DATE: 1994-02-10  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 2441  
; TYPE: PRT  
; ORGANISM: Mus  
; FEATURE:  
; NAME/KEY: VARIANT  
; LOCATION: (1)...(2441)  
; OTHER INFORMATION: Xaa = Any Amino Acid  
US-09-686-316-2

Query Match 68.2%; Score 88; DB 4; Length 2441;  
Best Local Similarity 66.7%; Pred. No. 3.3e-05;  
Matches 16; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 FRQPDVAVKLGPDYHKKIKQPM 24  
||||| :||| :|||  
Db 1112 FRQPDVQLLGIPDYFDIVKNPMD 1135

Search completed: July 27, 2004, 12:01:33  
Job time : 18 secs



GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 27, 2004, 12:00:41 ; Search time 43 Seconds  
(without alignments)  
174.784 Million cell updates/sec

Title: US-09-784-553C-19\_COPY\_29\_52  
Perfect score: 129  
Sequence: 1 FRQPDVAVKLGIPDYHKIIKQPM 24

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1288442 seqs, 313154207 residues

Total number of hits satisfying chosen parameters: 1288442

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*  
1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US03\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	129	100.0	111	12	US-09-784-553C-19
2	129	100.0	111	15	US-10-209-201C-19
3	129	100.0	801	14	US-10-146-473-42
4	113	87.6	111	12	US-09-784-553C-21
5	113	87.6	111	15	US-10-209-201C-21
6	110	85.3	111	12	US-09-784-553C-20
7	110	85.3	111	15	US-10-209-201C-20
8	110	85.3	140	9	US-09-764-864-1572
9	110	85.3	235	9	US-09-764-864-1159
10	107	82.9	1390	12	US-10-092-900A-224
11	96	74.4	947	14	US-10-293-822-1
12	93	72.1	111	12	US-09-784-553C-14
13	93	72.1	111	15	US-10-209-201C-14
14	92	71.3	714	16	US-10-437-963-156944
15	92	71.3	734	16	US-10-437-963-156947

Sequence 3827, Ap  
Sequence 64125, A  
Sequence 41, Appl  
Sequence 41, Appl  
Sequence 42, Appl  
Sequence 42, Appl  
Sequence 198928,  
Sequence 54, Appl  
Sequence 54, Appl  
Sequence 55, Appl  
Sequence 11, Appl  
Sequence 12, Appl  
Sequence 13, Appl  
Sequence 12, Appl  
Sequence 12, Appl  
Sequence 13, Appl  
Sequence 10, Appl  
Sequence 25, Appl  
Sequence 25, Appl  
Sequence 23, Appl  
Sequence 23, Appl  
Sequence 3800, A  
Sequence 63524, A  
Sequence 40087, A  
Sequence 244688,  
Sequence 154093, A  
Sequence 62363, A  
Sequence 198539,  
Sequence 172901,  
Sequence 172902,

16 91 70.5 723 15 US-10-369-493-3827  
17 90 69.8 346 12 US-10-425-114-64125  
18 89 69.0 61 12 US-09-839-479-41  
19 89 69.0 61 12 US-10-702-148-41  
20 89 69.0 61 12 US-10-376-537-42  
21 89 69.0 610 12 US-10-424-599-198828  
22 88 68.2 61 12 US-09-839-479-54  
23 88 68.2 61 12 US-10-702-148-54  
24 88 68.2 61 12 US-10-376-537-55  
25 88 68.2 112 12 US-09-784-553C-11  
26 88 68.2 112 12 US-09-784-553C-12  
27 88 68.2 112 12 US-09-784-553C-13  
28 88 68.2 112 15 US-10-209-201C-11  
29 88 68.2 112 15 US-10-209-201C-12  
30 88 68.2 112 15 US-10-209-201C-13  
31 88 68.2 2441 13 US-10-109-886-8  
32 88 68.2 2442 13 US-10-109-886-10  
33 87 67.4 113 12 US-09-784-553C-25  
34 87 67.4 113 15 US-10-209-201C-25  
35 86 66.7 113 12 US-09-784-553C-23  
36 86 66.7 113 15 US-10-209-201C-23  
37 86 66.7 343 12 US-10-425-114-38300  
38 86 66.7 575 12 US-10-425-114-63524  
39 85 65.9 96 12 US-10-425-114-40087  
40 85 65.9 128 12 US-10-424-599-244688  
41 85 65.9 655 16 US-10-437-963-154093  
42 81 62.8 478 12 US-10-425-114-62363  
43 81 62.8 563 16 US-10-437-963-198539  
44 81 62.8 751 16 US-10-437-963-172901  
45 81 62.8 791 16 US-10-437-963-172902

#### ALIGNMENTS

RESULT 1  
US-09-784-553C-19  
; Sequence 19, Application US/09784553C  
; Publication No. US20040043378A1  
; GENERAL INFORMATION:  
; APPLICANT: ZHOOU, MING-MING  
; TITLE OF INVENTION: METHODS OF IDENTIFYING MODULATORS OF BROMODOMAINS  
; FILE REFERENCE: 2459-1-003 CIP  
; CURRENT APPLICATION NUMBER: US/09/784,553C  
; CURRENT FILING DATE: 2003-07-08  
; PRIOR APPLICATION NUMBER: 09/510,314  
; PRIOR FILING DATE: 2000-02-22  
; NUMBER OF SEQ ID NOS: 63  
; SOFTWARE: Patent in version 3.0  
; SEQ ID NO 19  
; LENGTH: 111  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-784-553C-19

Query Match 100.0%; Score 129; DB 12; Length 111;  
Best Local Similarity 100.0%; Pred. No. 1.7e-12;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FRQPDVAVKLGIPDYHKIIKQPM 24  
DB 29 FRQPDVAVKLGIPDYHKIIKQPM 52

RESULT 2  
US-10-209-201C-19  
; Sequence 19, Application US/10209201C  
; Publication No. US20040009613A1  
; GENERAL INFORMATION:  
; APPLICANT: Verdin, Eric  
; APPLICANT: Bruland, Joan  
; APPLICANT: Ott, Melanie



Query Match 85.3%; Score 110; DB 12; Length 111;  
Best Local Similarity 83.3%; Pred. No. 1.7e-09;  
Matches 20; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 FROPVDVAVKLGLPDYHKIIKPMD 24  
DB 29 FYQPVDAIKNLDPYHKIIKPMD 52

## RESULT 7

US-10-209-201C-20  
; Sequence 20, Application US/10209201C  
; Publication No. US20040009613A1  
; GENERAL INFORMATION:  
; APPLICANT: Verdine, Eric  
; APPLICANT: Bruland, Joan  
; APPLICANT: Ott, Melanie  
; APPLICANT: Zhou, Ming-Ming  
; APPLICANT: Aggarwal, Aneel  
; TITLE OF INVENTION: Methods of Identifying Modulators of Bromodomains  
; FILE REFERENCE: 2459-1-003CIPDIV  
; CURRENT APPLICATION NUMBER: US/10/209,201C  
; CURRENT FILING DATE: 2002-07-31  
; PRIOR APPLICATION NUMBER: 09/784,553  
; PRIOR FILING DATE: 2001-02-16  
; PRIOR APPLICATION NUMBER: 09/510,314  
; PRIOR FILING DATE: 2000-02-22  
; NUMBER OF SEQ ID NOS: 60  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 20  
; LENGTH: 111  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-209-201C-20

Query Match 85.3%; Score 110; DB 15; Length 111;  
Best Local Similarity 83.3%; Pred. No. 1.7e-09;  
Matches 20; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 FROPVDVAVKLGLPDYHKIIKPMD 24  
DB 29 FYQPVDAIKNLDPYHKIIKPMD 52

## RESULT 8

US-09-764-864-1572  
; Sequence 1572, Application US/09764864  
; Patent No. US20020132753A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PT223  
; CURRENT APPLICATION NUMBER: US/09/764,864  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - consult PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 1792  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1572  
; LENGTH: 140  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-764-864-1572

Query Match 85.3%; Score 110; DB 9; Length 140;  
Best Local Similarity 83.3%; Pred. No. 2.3e-09;  
Matches 20; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 FROPVDVAVKLGLPDYHKIIKPMD 24  
DB 33 FYQPVDAIKNLDPYHKIIKPMD 56

## RESULT 9

US-09-764-864-1159  
; Sequence 1159, Application US/09764864  
; Patent No. US20020132753A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PT223  
; CURRENT APPLICATION NUMBER: US/09/764,864  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - consult PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 1792  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1159  
; LENGTH: 235  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (129)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (215)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (221)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-764-864-1159

Query Match 85.3%; Score 110; DB 9; Length 235;  
Best Local Similarity 83.3%; Pred. No. 4.1e-09;  
Matches 20; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 FROPVDVAVKLGLPDYHKIIKPMD 24  
DB 33 FYQPVDAIKNLDPYHKIIKPMD 56

## RESULT 10

US-10-092-900A-224  
; Sequence 224, Application US/10092900A  
; Publication No. US20040043382A1  
; GENERAL INFORMATION:  
; APPLICANT: Padigaru, Muralidhara  
; APPLICANT: Spytek, Kimberly A.  
; APPLICANT: Shenoy, Suresh G.  
; APPLICANT: Taupier Jr., Raymond J.  
; APPLICANT: Pena, Carol E.A.  
; APPLICANT: Li, Li  
; APPLICANT: Zernhuseen, Bryan D.  
; APPLICANT: Gusev, Vladimir Y.  
; APPLICANT: Ji, Weizhen  
; APPLICANT: Gorman, Linda  
; APPLICANT: Miller, Charles E.  
; APPLICANT: Kekuda, Ramesh  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Gangolli, Esha A.  
; APPLICANT: Vernet, Corine A.M.  
; APPLICANT: Guo, Xiaojia Sasha  
; APPLICANT: Tchernev, Velizar T.  
; APPLICANT: Fernandes, Elma R.  
; APPLICANT: Casman, Stacie J.  
; APPLICANT: Malyankar, Uriel M.  
; APPLICANT: Gerlach, Valerie  
; APPLICANT: Liu, Yi  
; APPLICANT: Anderson, David W.  
; APPLICANT: Spaderna, Steven K.  
; APPLICANT: Carterton, Elina  
; APPLICANT: Leite, Mario W.  
; APPLICANT: Zhong, Haihong  
; APPLICANT: Alsobrook, John P.  
; APPLICANT: Lepley, Denise M.  
; APPLICANT: Rieger, Daniel K.

```

; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: No. US20040043382A1e1 Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-290C
; CURRENT APPLICATION NUMBER: US/10/092,900A
; CURRENT FILING DATE: 2002-03-07
; PRIOR FILING DATE: USSN 60/274,322
; PRIOR FILING DATE: 2001-03-08
; PRIOR FILING DATE: USSN 60/283,675
; PRIOR FILING DATE: 2001-04-13
; PRIOR FILING DATE: USSN 60/338,092
; PRIOR FILING DATE: 2001-12-03
; PRIOR FILING DATE: USSN 60/274,281
; PRIOR FILING DATE: 2001-03-08
; PRIOR FILING DATE: USSN 60/274,191
; PRIOR FILING DATE: 2001-03-08
; PRIOR FILING DATE: USSN 60/325,681
; PRIOR FILING DATE: 2001-09-27
; PRIOR FILING DATE: USSN 60/304,354
; PRIOR FILING DATE: 2001-07-10
; PRIOR FILING DATE: USSN 60/279,995
; PRIOR FILING DATE: 2001-03-30
; PRIOR FILING DATE: USSN 60/294,899
; PRIOR FILING DATE: 2001-05-31
; PRIOR FILING DATE: USSN 60/287,424
; PRIOR FILING DATE: 2001-04-30
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 788
; SEQ ID NO 224
; LENGTH: 1390
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-092-900A-224

```

```

Query Match      82.9%; Score 107; DB 12; Length 1390;
Best Local Similarity 83.3%; Pred. No. 9.3e-08;
Matches 20; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 FRQPDVAVKLGDPYHKIKQPM 24
DB 83 FRQPDVAVKLNLPDYHKIKTPMD 106

```

```

RESULT 11
US-10-293-822-1
; Sequence 1, Application US/10293822
; Publication No. US20030083470A1
; GENERAL INFORMATION:
; APPLICANT: Jones, Michael H.
; TITLE OF INVENTION: TRANSCRIPTIONAL REGULATOR
; FILE REFERENCE: 06501-043001
; CURRENT APPLICATION NUMBER: US/10/293,822
; CURRENT FILING DATE: 2002-11-12
; PRIOR FILING DATE: 1999-10-15
; PRIOR FILING DATE: PCT/JP98/01782
; PRIOR FILING DATE: 1998-04-17
; PRIOR FILING DATE: JAPAN 9/116402
; PRIOR FILING DATE: 1997-04-18
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 947
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-822-1

```

```

Query Match      74.4%; Score 96; DB 14; Length 947;
Best Local Similarity 75.0%; Pred. No. 3.3e-06;
Matches 18; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 1 FRQPDVAVKLGDPYHKIKQPM 24
DB 52 FRQPDVAVKLNLPDYHKIKTPMD 75

```

```

RESULT 12
US-09-784-553C-14
; Sequence 14, Application US/09784553C
; Publication No. US20040043378A1
; GENERAL INFORMATION:
; APPLICANT: AGGARWAL, ANEEL
; APPLICANT: ZHOU, MING-MING
; TITLE OF INVENTION: METHODS OF IDENTIFYING MODULATORS OF BROMODOMAINS
; FILE REFERENCE: 2459-1-003 CIP
; CURRENT APPLICATION NUMBER: US/09/784,553C
; CURRENT FILING DATE: 2003-07-08
; PRIOR APPLICATION NUMBER: 09/510,314
; PRIOR FILING DATE: 2000-02-22
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 14
; LENGTH: 111
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-09-784-553C-14

```

```

Query Match      72.1%; Score 93; DB 12; Length 111;
Best Local Similarity 70.8%; Pred. No. 8.5e-07;
Matches 17; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

```

```

QY 1 FRQPDVAVKLGDPYHKIKQPM 24
DB 29 FRQPDVAVKLNLPDYHKIKRPM 52

```

```

RESULT 13
US-10-209-201C-14
; Sequence 14, Application US/10209201C
; Publication No. US20040009613A1
; GENERAL INFORMATION:
; APPLICANT: Verdin, Eric
; APPLICANT: Bruland, Joan
; APPLICANT: Ott, Melanie
; APPLICANT: Zhou, Ming-Ming
; APPLICANT: Aggarwal, Anneel
; TITLE OF INVENTION: Methods of Identifying Modulators of Bromodomains
; FILE REFERENCE: 2459-1-003CIPDIV
; CURRENT APPLICATION NUMBER: US/10/209,201C
; CURRENT FILING DATE: 2002-07-31
; PRIOR APPLICATION NUMBER: 09/784,553
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: 09/510,314
; PRIOR FILING DATE: 2000-02-22
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 14
; LENGTH: 111
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-10-209-201C-14

```

```

Query Match      72.1%; Score 93; DB 15; Length 111;
Best Local Similarity 70.8%; Pred. No. 8.5e-07;
Matches 17; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

```

```

QY 1 FRQPDVAVKLGDPYHKIKQPM 24
DB 29 FRQPDVAVKLNLPDYHKIKRPM 52

```

```

RESULT 14
US-10-437-963-156944
; Sequence 156944, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.

```

```
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 156944
; LENGTH: 714
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_56563C.1.pep
US-10-437-963-156944
```

```
Query Match 71.3%; Score 92; DB 16; Length 714;
Best Local Similarity 70.8%; Pred. No. 1e-05;
Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 FROPVDAVKLGLPDYHKIKQPM 24
Db 187 FDSFVDAVKLNIPIFYFOIKKPM 210
```

```
RESULT 15
US-10-437-963-156947
; Sequence 156947, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 156947
; LENGTH: 734
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_56566C.1.pep
US-10-437-963-156947
```

```
Query Match 71.3%; Score 92; DB 16; Length 734;
Best Local Similarity 70.8%; Pred. No. 1.1e-05;
Matches 17; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
```

```
QY 1 FROPVDAVKLGLPDYHKIKQPM 24
Db 181 FDSFVDAVKLNIPIFYFOIKKPM 204
```

```
Search completed: July 27, 2004, 12:06:09
Job time : 44 secs
```





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

CM protein - protein search, using sw model

Run on: July 27, 2004, 12:05:22 ; Search time 16 Seconds  
(without alignments)  
144.287 Million cell updates/sec

Title: US-09-784-553C-19\_COPY\_29\_52  
Perfect score: 129  
Sequence: 1 FRQPDVAVKGLPDYHKIIKQPM 24

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 78:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	129	100.0	733	2 T28145	RING3 kinase - chicken
2	129	100.0	754	2 A56619	female sterile hom
3	113	87.6	2038	2 A43742	female sterile hom
4	97	75.2	374	2 T33328	hypothetical prote
5	93	72.1	2027	2 S60123	hypothetical prote
6	93	72.1	2036	2 G85564	protein R10E11.1 [
7	91	70.5	461	2 D96757	hypothetical prote
8	88	68.2	2414	2 A54277	transcription adap
9	88	68.2	2440	2 S39162	transcription coac
10	88	68.2	2441	2 S39161	CREB-binding prote
11	86	66.7	440	2 H86312	F2HJ5-2 protein -
12	86	66.7	1087	2 T22847	hypothetical prote
13	86	66.7	1250	2 T22845	hypothetical prote
14	86	66.7	3190	2 T13828	CREB-binding prote
15	83	64.3	678	2 T49984	bromodomain protei
16	78	60.5	400	2 T00472	probable RING3 pro
17	77	59.7	578	2 T40984	transcription fact
18	77	59.7	638	2 S67605	hypothetical prote
19	76	58.9	686	2 S55955	bromodomain protei
20	73	56.6	361	2 T42517	bromodomain protei
21	72	55.8	703	2 T48600	kinase-like protei
22	72	55.8	769	2 E96613	hypothetical prote
23	70	54.3	766	2 A86198	hypothetical prote
24	69	53.5	1051	2 S55259	TIF1 protein - mou
25	67	51.9	405	2 T21433	hypothetical prote
26	67	51.9	452	2 T21435	hypothetical prote
27	67	51.9	510	2 T21430	hypothetical prote
28	63	48.8	369	2 T46098	hypothetical prote
29	63	48.8	586	2 T47620	histon acetyltrans

ALIGNMENTS

RESULT 1

T28145  
RING3 kinase - chicken  
C:Species: Gallus gallus (chicken)  
C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jun-2000  
C:Accession: T28145  
R:Milne S.; Kaufman, J.; Beck, S.  
submitted to the EMBL Data Library, May 1998  
A:Description: DNA sequencing and analysis of the chicken major histocompatibility comp  
A:Reference number: Z20475  
A:Accession: T28145  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-733 <ML>  
A:Cross-references: EMBL:AL023516; NID:e1292539; PID:e1292549; PIDN:CAA18965.1  
A:Experimental source: clone cH12  
C:Genetics:  
A:Gene: RING3  
A:Map position: 16  
A:Introns: 64/3; 110/3; 158/1; 227/3; 351/3; 394/3; 479/3; 546/2; 650/1; 691/1  
C:Superfamily: unassigned bromodomain proteins; bromodomain homology  
F:52-103/Domain: bromodomain homology <BRO1>  
F:323-380/Domain: bromodomain homology <BRO2>

Query Match 100.0%; Score 129; DB 2; Length 733;  
Best Local Similarity 100.0%; Pred. No. 2.3e-12;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 FRQPDVAVKGLPDYHKIIKQPM 24  
DB 52 FRQPDVAVKGLPDYHKIIKQPM 75  
RESULT 2  
A56619  
female sterile homeotic (fsh) homolog RING3 - human  
C:Species: Homo sapiens (man)  
C>Date: 21-Jul-1995 #sequence\_revision 03-Aug-1995 #text\_change 20-Sep-1999  
C:Accession: A56619; S18860; S40781  
R:Beck, S.; Hanson, I.; Kelly, A.; Pappin, D.J.; Trowsdale, J.  
DNA Seq. 2, 203-210, 1992  
A:Title: A homologue of the Drosophila female sterile homeotic (fsh) gene in the class I  
A:Reference number: A56619; MUID:92329974; PMID:1352711  
A:Accession: A56619  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-754 <BRC>  
A:Cross-references: EMBL:X62083; NID:g31471; PIDN:AAA68890.1; PID:g31472; EMBL:M80613; I  
A>Note: sequence extracted from NCBI backbone (NCBIP:108781)  
C:Genetics:  
A:Gene: RING3  
C:Superfamily: unassigned bromodomain proteins; bromodomain homology

C;Keywords: duplication  
 F:52-109/Domain: bromodomain homology <BRO1>  
 F:325-382/Domain: bromodomain homology <BRO2>

Query Match 100.0%; Score 129; DB 2; Length 754;  
 Best Local Similarity 100.0%; Pred. No. 2.4e-12;  
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FRQPVDAVKGGLPDYHKIKQPM 24  
 ||||| ||||| ||||| ||||| |||||  
 DB 52 FRQPVDAVKGGLPDYHKIKQPM 75

## RESULT 3

A43742 female sterile homeotic protein, 205K - fruit fly (*Drosophila melanogaster*)  
 N;Alternate names: membrane protein fsh, 205K  
 N;Contains: female sterile homeotic protein, 110K  
 C;Species: *Drosophila melanogaster*  
 C;Date: 03-Mar-1993 #sequence\_revision 03-Mar-1993 #text\_change 20-Sep-1999  
 C;Accession: A43742; B43742  
 R;Haynes, S.R.; Mozer, B.A.; Bhatia-Dey, N.; Dawid, I.B.  
 Dev. Biol. 134, 246-257, 1989  
 A;Title: The *Drosophila* fsh locus, a maternal effect homeotic gene, encodes apparent mem  
 A;Reference number: A43742; MUID:89276730; PMID:2567251  
 A;Accession: A43742  
 A;Status: preliminary  
 A;Molecule type: mRNA  
 A;Residues: 1-2038 <HAY>  
 A;Cross-references: EMBL:M23221; NID:g157452; PIDN:AAA28540.1; PID:g157453  
 A;Accession: B43742  
 A;Status: preliminary  
 A;Molecule type: mRNA  
 A;Residues: 1-1106 <HA2>  
 A;Cross-references: EMBL:M23222  
 C;Genetics:  
 A;Gene: fsh  
 A;Cross-references: FlyBase:FBgn004656  
 C;Superfamily: unassigned bromodomain proteins; bromodomain homology  
 C;Keywords: alternative splicing; transmembrane protein  
 F:1-2038/Product: female sterile homeotic protein, 205K #status predicted <WA2>  
 F:1-1106/Product: female sterile homeotic protein, 110K #status predicted <WA2>  
 F:59-116/Domain: bromodomain homology <BRO1>  
 F:503-560/Domain: bromodomain homology <BRO2>

Query Match 87.6%; Score 113; DB 2; Length 2038;  
 Best Local Similarity 87.6%; Pred. No. 2.9e-09;  
 Matches 21; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 FRQPVDAVKGGLPDYHKIKQPM 24  
 ||||| ||||| ||||| ||||| |||||  
 DB 59 FQPVDAKLNLPDYHKIKQPM 82

## RESULT 4

T33328 hypothetical protein Fl3C5.2 - *Caenorhabditis elegans*  
 C;Species: *Caenorhabditis elegans*  
 C;Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 02-Sep-2000  
 C;Accession: T33328  
 R;Wohlmann, P.; Hawkins, J.; Gillam, B.  
 Submitted to the EMBL Data Library, July 1998  
 A;Description: The sequence of *C. elegans* cosmid F13C5.  
 A;Reference number: Z21324  
 A;Accession: T33328  
 A;Status: preliminary; translated from GB/EMBL/DDBJ  
 A;Molecule type: DNA  
 A;Residues: 1-374 <WOH>  
 A;Cross-references: EMBL:AF077531; PIDN:AAC64610.1; GSPDB:GN00028; CESP:F13C5.2  
 A;Experimental source: strain Bristol N2; clone F13C5  
 C;Genetics:  
 A;Gene: CESP:F13C5.2  
 A;Map position: X

A;introns: 25/3; 135/3; 189/3; 313/1  
 C;Superfamily: bromodomain homology  
 F:141-198/Domain: bromodomain homology <BRO>

Query Match 75.2%; Score 97; DB 2; Length 374;  
 Best Local Similarity 70.8%; Pred. No. 1.4e-07;  
 Matches 17; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 FRQPVDAVKGGLPDYHKIKQPM 24  
 ||||| ||||| ||||| ||||| |||||  
 DB 141 FRQPVDAVKGGLPDYHKIKQPM 164

## RESULT 5

S60123 hypothetical protein R10E11.1 - *Caenorhabditis elegans*  
 C;Species: *Caenorhabditis elegans*  
 C;Date: 13-Jan-1996 #sequence\_revision 01-Mar-1996 #text\_change 02-Aug-2002  
 C;Accession: S60123; S40713  
 R;Ainscough, R.; Mortimore, B.  
 Submitted to the EMBL Data Library, November 1995  
 A;Reference number: S60123  
 A;Accession: S60123  
 A;Molecule type: DNA  
 A;Residues: 1-2027 <AIN>  
 A;Cross-references: EMBL:Z29095; NID:g436453; PID:g1067032  
 A;Note: this is a revision to the sequence from reference S40713  
 R;Ainscough, R.; Mortimore, B.  
 Submitted to the EMBL Data Library, December 1993  
 A;Reference number: S40713  
 A;Accession: S40713  
 A;Molecule type: DNA  
 A;Residues: 1-466; 'CRYITRRVASFSLGK', 467, 'FEHR', 474-475, 'KRLFPKISHSHF', 479-1986, 'GQ'  
 A;Cross-references: EMBL:Z29095  
 A;Note: this sequence has been revised in reference S60123  
 C;Genetics:  
 A;introns: 14/1; 39/3; 302/3; 424/3; 467/1; 517/1; 688/1; 1759/1; 1828/2; 1892/3; 1964/3  
 C;Superfamily: transcription coactivator CREB-binding protein; bromodomain homology  
 F:889-946/Domain: bromodomain homology <BRO>

Query Match 72.1%; Score 93; DB 2; Length 2027;  
 Best Local Similarity 70.8%; Pred. No. 4.6e-06;  
 Matches 17; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 FRQPVDAVKGGLPDYHKIKQPM 24  
 ||||| ||||| ||||| ||||| |||||  
 DB 889 FRQPVDAKLNLPDYHEIKRPM 912

## RESULT 6

G88564 protein R10E11.1 [imported] - *Caenorhabditis elegans*  
 C;Species: *Caenorhabditis elegans*  
 C;Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 02-Aug-2002  
 C;Accession: G88564  
 R;anonymus, The C. elegans Sequencing Consortium.  
 Science 282, 2012-2018, 1998  
 A;Title: Genome sequence of the nematode *C. elegans*: a platform for investigating biolog  
 A;Reference number: A75000; MUID:99069613; PMID:9851916  
 A;Note: see websites genome.wustl.edu/gsc/C\_elegans/ and www.sanger.ac.uk/Projects/C\_eleg  
 A;Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and  
 A;Accession: G88564  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-2056 <STO>  
 A;Cross-references: GB:chr\_III; PIDN:CAA82353.1; PID:g3979836; GSPDB:GN00021; CESP:R10E11  
 C;Genetics:  
 A;Gene: R10E11.1  
 A;Map position: 3  
 C;Superfamily: transcription coactivator CREB-binding protein; bromodomain homology  
 Query Match 72.1%; Score 93; DB 2; Length 2056;  
 Best Local Similarity 70.8%; Pred. No. 4.7e-06;

Matches 17; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 FRQPVDAVKLGDPDYHKIIKQPM 24  
 Db 889 FRVFDVAKULNPDPYHIIKRPMD 912

RESULT 7  
 D96757  
 hypothetical protein T18K17.19 [imported] - Arabidopsis thaliana  
 C/Species: Arabidopsis thaliana (mouse-ear cress)  
 C/Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
 C/Accession: D96757  
 R/Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
 ansen, N.F.; Hughes, B.; Huizar, L.  
 Nature 408, 816-820, 2000  
 A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Mailfi, R.; Marziali,  
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
 A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
 A/Reference number: A86141; MUID:21016719; PMID:11130712  
 A/Accession: D96757  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-461 <STO>  
 A/Cross-references: GB:AE005173; NID:g6598866; PIDN:AAF18720.1; GSPDB:GN00141  
 C/Genetics:  
 A/Gene: T18K17.19  
 A/Map position: 1

Query Match 70.5%; Score 91; DB 2; Length 461;  
 Best Local Similarity 70.8%; Pred. No. 1.7e-06;  
 Matches 17; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 FRQPVDAVKLGDPDYHKIIKQPM 24  
 Db 139 FNTFVDVTLGLHDYHNIKRPMD 162

RESULT 8  
 A54277  
 transcription adaptor protein p300 - human  
 C/Species: Homo sapiens (man)  
 C/Date: 09-Sep-1994 #sequence\_revision 09-Sep-1994 #text\_change 02-Aug-2002  
 C/Accession: A54277; S60344  
 R/Eckner, R.; Ewen, M.E.; Newsome, D.; Gerdes, M.; DeCaprio, J.A.; Lawrence, J.B.; Livin  
 Genes Dev. 8, 869-884, 1994  
 A/Title: Molecular cloning and functional analysis of the adenovirus E1A-associated 300-  
 A/Reference number: A54277; MUID:95011587; PMID:7523245  
 A/Accession: A54277  
 A/Status: preliminary  
 A/Molecule type: mRNA  
 A/Residues: 1-2414 <ECK>  
 A/Cross-references: GB:U01877; NID:g495300; PIDN:AAAI8639.1; PID:g495301  
 A/Note: In the authors' translation 941-Ser is shown after 961 and consequently, residue  
 R/Lundblad, J.R.; Kwok, R.P.S.; Laurance, M.E.; Harter, M.L.; Goodman, R.H.  
 Nature 374, 85-88, 1995  
 A/Title: Adenoviral E1A-associated protein p300 as a functional homologue of the transcr  
 A/Reference number: S60344; MUID:95174889; PMID:7870179  
 A/Accession: S60344  
 A/Status: preliminary  
 A/Molecule type: protein  
 A/Residues: 552-660 <LUN>  
 C/Genetics:  
 A/Gene: GDB:EF300  
 A/Cross-references: GDB:9862958; OMIM:502700  
 A/Map position: 22q13.2-22q13.2  
 C/Superfamily: transcription coactivator CREB-binding protein; bromodomain homology  
 C/Keywords: phosphoprotein; transcription; zinc finger  
 F:1075-1132/Domain: bromodomain homology <BRO>

F:89.507,1136,1295,1497,1834,1977,2062,2320/Binding site: phosphate (Ser) (covalent) (by  
 F:1734/Binding site: phosphate (Ser) (covalent) (by protein kinase A) #status predicted

Query Match 68.2%; Score 88; DB 2; Length 2414;  
 Best Local Similarity 66.7%; Pred. No. 3.6e-05;  
 Matches 16; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 FRQPVDAVKLGDPDYHKIIKQPM 24  
 Db 1075 FRQVDPQLGIPDYFDIVKSPMD 1098

RESULT 9  
 S39162  
 transcription coactivator CREB-binding protein - human  
 N/Alternate names: CBP; RSTS; Rubinstein-Taybi syndrome (RTS) protein  
 C/Species: Homo sapiens (man)  
 C/Date: 07-Oct-1994 #sequence\_revision 17-Nov-1995 #text\_change 02-Aug-2002  
 C/Accession: S39162; S60345; I58096  
 R/Chrivia, J.C.; Kwok, R.P.S.; Lamb, N.; Hagiwara, M.; Montminy, M.R.; Goodman, R.H.  
 Nature 365, 855-859, 1993  
 A/Title: Phosphorylated CREB binds specifically to the nuclear protein CBP.  
 A/Reference number: S39161; MUID:94019866; PMID:8413673  
 A/Accession: S39162  
 A/Molecule type: mRNA  
 A/Residues: 1-2440 <CHR>  
 A/Note: differences with the mouse sequence are shown  
 R/Lundblad, J.R.; Kwok, R.P.S.; Laurance, M.E.; Harter, M.L.; Goodman, R.H.  
 Nature 374, 85-88, 1995  
 A/Title: Adenoviral E1A-associated protein p300 as a functional homologue of the transcr  
 A/Reference number: S60344; MUID:95174889; PMID:7870179  
 A/Accession: S60345  
 A/Status: preliminary  
 A/Molecule type: protein  
 A/Residues: 'S', 574-670, 'T', 672-681 <LUN>  
 R/Petrij, F.; Giles, R.H.; Dauwerse, H.G.; Saris, J.J.; Hennekam, R.C.; Maasuno, M.; Tom  
 Nature 376, 348-351, 1995  
 A/Title: Rubinstein-Taybi syndrome caused by mutations in the transcriptional co-activ  
 A/Reference number: I58096; MUID:95356817; PMID:7630403  
 A/Accession: I58096  
 A/Status: translation not shown; translated from GB/EMBL/DBJ  
 A/Molecule type: mRNA  
 A/Residues: 352-356 <PET>  
 A/Cross-references: GB:U89354; NID:gi1888537; GB:S78936; NID:gl041931  
 A/Note: This translation is not annotated in GenBank entry S78936, release 112.0  
 A/Note: This sequence with a termination mutation is from a patient with Rubinstein-Tay  
 C/Genetics:  
 A/Gene: GDB:CREBBP; RTS; CBP; RSTS  
 A/Cross-references: GDB:437159; OMIM:180849; OMIM:600140  
 A/Map position: 16p13.3-16p13.3  
 A/Note: defects in this gene may result in Rubinstein-Taybi syndrome  
 C/Superfamily: transcription coactivator CREB-binding protein; bromodomain homology  
 C/Keywords: phosphoprotein; transcription; zinc finger  
 F:462-661/Domain: CREB binding #status predicted <CBB>  
 F:1111-1168/Domain: bromodomain homology <BRO>  
 F:1283-1311/Region: zinc finger CCCC motif  
 F:1707-1732/Region: zinc finger CCCC motif  
 F:78.381,745,1172,1533,2063,2354/Binding site: phosphate (Ser) (covalent) (by calmodulin  
 F:1771/Binding site: phosphate (Ser) (covalent) (by protein kinase A) #status predicted

Query Match 65.2%; Score 88; DB 2; Length 2440;  
 Best Local Similarity 66.7%; Pred. No. 3.6e-05;  
 Matches 16; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 FRQPVDAVKLGDPDYHKIIKQPM 24  
 Db 1111 FRQVDPQLGIPDYFDIVKSPMD 1134

RESULT 10  
 S39161  
 CREB-binding protein - mouse  
 C/Species: Mus musculus (house mouse)

C;Date: 25-Feb-1994 #sequence\_revision 10-Nov-1995 #text\_change 02-Aug-2002  
 C;Accession: S39161  
 R;Christia, J.C.; Kwok, R.P.S.; Lamb, N.; Hagiwara, M.; Montminy, M.R.; Goodman, R.H.  
 Nature 365, 855-859, 1993  
 A;Title: Phosphorylated CREB binds specifically to the nuclear protein CBP.  
 A;Reference number: S39161; MUID:94019866; PMID:8413673  
 A;Accession: S39161  
 A;Status: preliminary  
 A;Molecule type: mRNA  
 A;Residues: 1-2441 <CHR>  
 A;Cross-references: GB:S66385; NID:9435854; PIDN:AAB28651.1; PID:9435855  
 C;Superfamily: transcription coactivator CREB-binding protein; bromodomain homology  
 F;1112-1169/Domain: bromodomain homology <BRO>

Query Match 58.2%; Score 88; DB 2; Length 2441;  
 Best Local Similarity 56.7%; Pred. No. 3.6e-05;  
 Matches 16; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 FRQPVDAVKLGLPDYHKIKQPM 24  
 ||||| ||||| ||||| ||||| |||||  
 Db 1112 FRQVPDQLGIPDYFDVKNPMD 1135

RESULT 11  
 H86312  
 F2H15.2 protein - Arabidopsis thaliana  
 C;Species: Arabidopsis thaliana (mouse-ear cress)  
 C;Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 02-Nov-2001  
 C;Accession: H86312  
 R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
 Jensen, N.F.; Hughes, B.; Ruizar, L.  
 Nature 408, 816-820, 2000  
 A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,  
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
 A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
 Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
 A;Reference number: A86141; MUID:21016719; PMID:11130712  
 A;Accession: H86312  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-440 <STO>  
 A;Cross-references: GB:AE005172; NID:99665057; PIDN:AAF97259.1; GSPDB:GNC00141  
 C;Genetics:  
 A;Map position: 1

Query Match 66.7%; Score 86; DB 2; Length 440;  
 Best Local Similarity 66.7%; Pred. No. 1e-05;  
 Matches 15; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 FRQPVDAVKLGLPDYHKIKQPM 24  
 ||||| ||||| ||||| ||||| |||||  
 Db 152 FNVPVDAKGLGLDHYHNIVKPEMD 175

RESULT 12  
 T22847  
 hypothetical protein F57C7.1b - Caenorhabditis elegans  
 C;Species: Caenorhabditis elegans  
 C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 02-Sep-2000  
 C;Accession: T22847  
 R;White, S.  
 Submitted to the EMBL Data Library, February 1996  
 A;Reference number: Z19625  
 A;Accession: T22847  
 A;Status: preliminary; translated from GB/EMBL/DDBJ  
 A;Molecule type: DNA  
 A;Residues: 1-1087 <WIL>  
 A;Cross-references: EMBL:Z69646; PIDN:CAA93475.1; GSPDB:GNC00028; CESP:F57C7.1b  
 A;Experimental source: clone F57C7  
 C;Genetics:

A;Gene: CESP:F57C7.1b  
 A;Map position: X  
 A;Introns: 262/3; 351/3; 391/3; 627/3; 795/2; 957/1  
 C;Superfamily: bromodomain homology  
 F;307-364/Domain: bromodomain homology <BRO1>  
 F;579-636/Domain: bromodomain homology <BRO2>

Query Match 66.7%; Score 86; DB 2; Length 1087;  
 Best Local Similarity 58.3%; Pred. No. 2.9e-05;  
 Matches 14; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 FRQPVDAVKLGLPDYHKIKQPM 24  
 ||||| ||||| ||||| ||||| |||||  
 Db 307 FQLEPDAIKLEIPEYHNIVNTPMD 330

RESULT 13  
 T22845  
 hypothetical protein F57C7.1a - Caenorhabditis elegans  
 C;Species: Caenorhabditis elegans  
 C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 02-Sep-2000  
 C;Accession: T22845  
 R;White, S.  
 Submitted to the EMBL Data Library, February 1996  
 A;Reference number: Z19625  
 A;Accession: T22845  
 A;Status: preliminary; translated from GB/EMBL/DDBJ  
 A;Molecule type: DNA  
 A;Residues: 1-1250 <WIL>  
 A;Cross-references: EMBL:Z69646; PIDN:CAA93473.1; GSPDB:GNC00028; CESP:F57C7.1a  
 A;Experimental source: clone F57C7  
 C;Genetics:  
 A;Gene: CESP:F57C7.1a  
 A;Map position: X  
 A;Introns: 262/3; 351/3; 391/3; 431/3; 667/3; 835/2; 997/1; 1127/1; 1218/3  
 C;Superfamily: bromodomain homology  
 F;307-364/Domain: bromodomain homology <BRO1>  
 F;619-676/Domain: bromodomain homology <BRO2>

Query Match 66.7%; Score 86; DB 2; Length 1250;  
 Best Local Similarity 59.3%; Pred. No. 3.5e-05;  
 Matches 14; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 FRQPVDAVKLGLPDYHKIKQPM 24  
 ||||| ||||| ||||| ||||| |||||  
 Db 307 FQLEPDAIKLEIPEYHNIVNTPMD 330

RESULT 14  
 T13828  
 CREB-binding protein homolog - fruit fly (Drosophila melanogaster)  
 C;Species: Drosophila melanogaster  
 C;Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 17-Nov-2000  
 C;Accession: T13828  
 R;Akimaru, H.; Chen, Y.; Dai, P.; Hou, D.X.; Nonaka, M.; Smolik, S.M.; Armstrong, S.; Go  
 Nature 386, 735-738, 1997  
 A;Title: Drosophila CBP is a co-activator of cubitus interruptus in hedgehog signalling.  
 A;Reference number: Z17785; MUID:97263578; PMID:9109493  
 A;Accession: T13828  
 A;Status: preliminary; translated from GB/EMBL/DDBJ  
 A;Molecule type: mRNA  
 A;Residues: 1-3190 <AKI>  
 A;Cross-references: EMBL:U88570; NID:g1916929; PID:g1916930; PIDN:AAB53050.1  
 C;Genetics:  
 A;Cross-references: FlyBase:FBgn0015624  
 A;Map position: X  
 C;Superfamily: bromodomain homology  
 F;1723-1780/Domain: bromodomain homology <BRO>

Query Match 66.7%; Score 86; DB 2; Length 3190;  
 Best Local Similarity 62.5%; Pred. No. 0.0001;  
 Matches 15; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

```
QY      1  FROPVDAVKLGLEDYHKIKQPM 24
      |||||  |||||  |||||  |||||
Db      1723  FRYPVDQALGIDYFEIVKPKND 1745

RESULT 15
T49984
bromodomain protein-like - Arabidopsis thaliana
N;Alternate names: protein F12B17.100
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Sep-2000
C;Accession: T49984
R;Bevan, M.; Bancroft, I.; Mewes, H.W.; Rudd, S.; Lemcke, K.; Mayer, K.F.X.
submitted to the Protein Sequence Database, April 2000
A;Reference number: Z25026
A;Accession: T49984
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-678 <BEV>
A;Cross-references: EMBL:AL353995; GSPDB:GN00063; ATSP:F12B17.100
A;Experimental source: cultivar Columbia; BAC clone F12B17
C;Genetics:
A;Gene: ATSP:F12B17.100
A;Map position: 5
A;Introns: 73/3; 560/3; 629/1; 656/2
C;Superfamily: bromodomain homology
F;269-326/Domain: bromodomain homology <BRO>

Query Match      64.3%; Score 83; DB 2; Length 678;
Best Local Similarity 62.5%; Pred. No. 5,1e-05;
Matches 15; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY      1  FROPVDAVKLGLEDYHKIKQPM 24
      |||||  |||||  |||||  |||||
Db      269  FLNPDVVGVLGLEDYHRIVDKPM 292

Search completed: July 27, 2004, 12:08:51
Job time : 17 secs
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 27, 2004, 12:01:36 ; Search time 14 Seconds  
(without alignments)  
89.263 Million cell updates/sec

Title: US-09-784-553C-19\_COPY\_29\_52  
Perfect score: 129  
Sequence: 1 FRQPVDAVKLGLPDYHKIKQPMQ 24

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	129	100.0	801	1 BRD2 HUMAN	P25440 homo sapien
2	113	87.6	2038	1 FSH DROME	P13709 drosophila
3	110	85.3	726	1 BRD3 HUMAN	Q15059 homo sapien
4	107	82.9	1362	1 BRD4 HUMAN	O60885 homo sapien
5	93	72.1	2056	1 CEP1 CAEL	P34545 caenorhabdi
6	88	68.2	2414	1 P300 HUMAN	Q09472 homo sapien
7	88	68.2	2441	1 CBP MOUSE	P45481 mus musculus
8	88	68.2	2442	1 CBP HUMAN	Q92793 homo sapien
9	79	61.2	727	1 YK82 SCHPO	Q9hgq4 schizosacch
10	76	58.9	686	1 BDF1 YEAST	P35817 saccharomyc
11	69	53.5	1050	1 TFI1 HUMAN	O15164 homo sapien
12	69	53.5	1051	1 TFI1A MOUSE	O64127 mus musculus
13	68	52.7	2269	1 WDR3 HUMAN	Q9ns16 homo sapien
14	67.5	52.3	1127	1 TFIIG HUMAN	Q9upn9 homo sapien
15	61	47.3	439	1 GCN5 YEAST	Q03330 saccharomyc
16	58	45.0	1872	1 T2D1 HUMAN	P21675 homo sapien
17	57	44.2	2065	1 T2D1 DROME	P51123 drosophila
18	55	42.6	1484	1 CES2 HUMAN	Q9bxf3 homo sapien
19	55	42.6	1638	1 BRM DROME	P25439 drosophila
20	52	40.3	263	1 PCNT ARATH	Q9mgq7 arabidopsis
21	52	40.3	283	1 PCNA BRANA	Q43124 brassica na
22	52	40.3	264	1 PCN2 ARATH	Q9zw53 arabidopsis
23	52	40.3	1972	1 BA2B HUMAN	Q9uif8 homo sapien
24	52	40.3	2130	1 BA2B CHICK	Q9del13 gallus gall
25	50	38.8	236	1 PCNA SOYBN	P22177 glycine max
26	50	38.8	264	1 PCNA POPNI	Q9may3 populus nig
27	49	38.0	739	1 PURL LISIN	Q92an9 listeria in
28	49	38.0	739	1 PURL LISMO	Q8y6c1 listeria mo
29	49	38.0	830	1 GCL2 MOUSE	Q9jhd2 mus musculus
30	49	38.0	837	1 GCL2 HUMAN	Q92830 homo sapien
31	49	38.0	1058	1 BRD1 HUMAN	O95696 homo sapien
32	49	38.0	1556	1 BA1A HUMAN	Q9nrl2 h bromodoma
33	48.5	37.6	199	1 YECD_ECOLI	P37347 escherichia

## ALIGNMENTS

## RESULT 1

ID	BRD2 HUMAN	STANDARD;	PRT;	801 AA.
AC	P25440; Q00699; Q00700; Q15310; Q969U4;			
DT	01-MAY-1992 (Rel. 22, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Bromodomain-containing protein 2 (RING3 protein) (O27.1.1).			
GN	BRD2 OR RING3 OR KIAA9001.			
OS	Homo sapiens (Human)			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]_TaxID=9606;			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=T-cell;			
RX	MEDLINE=92329974; PubMed=1352711;			
RA	Beck S., Hanson I., Kelly A., Pappin D.J.C., Tröwsdale J.,			
RT	"A homologue of the Drosophila female sterile homeotic (fsh) gene in			
RT	the class II region of the human MHC."			
RL	DNA Seq. 2:203-210(1992).			
RN	[2]			
RP	REVISIONS TO N-TERMINUS.			
RX	MEDLINE=96376536; PubMed=8781126;			
RA	Thorpe K.L., Abdulla S., Kaufman J., Trowsdale J., Beck S.;			
RT	"Phylogeny and structure of the RING3 gene."			
RL	Immunogenetics 44:391-396(1996).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Bone marrow;			
RA	Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayashi Y.,			
RA	Sato S., Nagase T., Seki T., Ishikawa K.-I., Tabata S.;			
RL	Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RA	Thorpe K.;			
RL	Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.			
CC	-!- SUBCELLULAR LOCATION: Nuclear (Potential).			
CC	-!- SIMILARITY: Contains 2 bromodomains.			
CC	-!- SIMILARITY: Contains 1 ET domain.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>			
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	-----			
CC	EMBL; X62083; CAA43996.1; -			
DR	EMBL; M80613; AAA68890.1; -			
DR	EMBL; D42040; BAA07641.1; -			
DR	EMBL; Z84497; CAC69991.1; -			
DR	EMBL; Z96104; CAC69989.1; -			
DR	EMBL; X96670; CAA65450.1; -			
DR	HSSP; Q92831; 1991.			

34	48	37.2	542	1	BRD1 SCHPO	Q09948 schizosacch
35	48	37.2	1214	1	BRF3 HUMAN	Q9ulid4 homo sapien
36	48	37.2	1332	1	SPT7 YEAST	P35177 saccharomyc
37	47	36.4	238	1	RIBB ARCFU	O28173 archaeoglob
38	47	36.4	264	1	PCNA DAUCA	Q00268 daucus caro
39	47	36.4	264	1	PCNA TOBAC	O82797 nicotiana t
40	47	36.4	268	1	PCNA CATRO	P24314 catharanthu
41	47	36.4	1196	1	DNBI HSV11	P04296 herpes simp
42	47	36.4	1196	1	DNBI HSV1F	P17469 herpes simp
43	47	36.4	1196	1	DNBI HSV1K	P17470 herpes simp
44	47	36.4	1479	1	BA1B MOUSE	Q92277 mus musculu
45	46.5	36.0	733	1	YC40 HUMAN	Q9ulid0 homo sapien

```

EMBL; M15763; AAA70423.1; -
EMBL; M15764; AAA70432.1; -
PIR; A43742; A43742.
HSSP; Q92831; 1B91.
FlyBase; FBgn004656; fs(1)h.
InterPro; IPR001487; Bromodomain.
Pfam; PF00439; bromodomain; 2.
PRINTS; PR00503; BROMODOMAIN.
SMART; SMC0297; BROMO; 2.
PROSITE; PS00633; BROMODOMAIN_1; 2.
PROSITE; PS50014; BROMODOMAIN_2; 2.
Developmental protein; Bromodomain; Transmembrane; Repeat.
DOMAIN 51 123 BROMODOMAIN 1.
DOMAIN 495 567 BROMODOMAIN 2.
DOMAIN 945 1106 ET DOMAIN.
TRANSMEM 330 350 POTENTIAL.
TRANSMEM 451 471 POTENTIAL.
TRANSMEM 750 770 POTENTIAL.
TRANSMEM 790 810 POTENTIAL.
TRANSMEM 816 830 POTENTIAL.
TRANSMEM 874 894 POTENTIAL.
TRANSMEM 1731 1751 POTENTIAL.
TRANSMEM 1939 1959 POTENTIAL.
VARIANT 909 G -> A.
VARIANT 1022 1022 H -> RKPYY.
SEQUENCE 2038 AA; 205332 MW; 849E0706D50A0098 CRC64;

Query Match 87.6%; Score 113; DB 1; Length 2038;
Best Local Similarity 87.5%; Pred. No. 1.2e-09;
Matches 21; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 FROPVDAVKLGLPDYHKIIKQPM 24
:|||||
DDB 59 FQOPVDAKKLNPDYHKIIKQPM 82

RESULT 3
BRD3_HUMAN
ID BRD3_HUMAN STANDARD; PRT; 726 AA.
Q15059; Q92645;
16-OCT-2001 (Rel. 40, Created)
16-OCT-2001 (Rel. 40, Last sequence update)
10-OCT-2003 (Rel. 42, Last annotation update)
Bromodomain-containing protein 3 (RING3-like protein).
BRD3 OR RING3L OR KIAA0043.
Brom sapientis (Human).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_taxid=9606;
[1]
SEQUENCE FROM N.A.
TISUE=Bone marrow;
MEDLINE=96051398; PubMed=7584044;
Nomura N., Nagase T., Miyajima N., Suzuki T., Tanaka A., Sato S.,
Seki N., Kawababayashi Y., Ishikawa K.-I., Tabata S.;
"Prediction of the coding sequences of unidentified human genes. II.
The coding sequences of 40 new genes (KIAA0041-KIAA0080) deduced by
analysis of cDNA clones from human cell line KG-1.";
DNA Res. 1:223-229 (1994).
[2]
SEQUENCE OF 363-726 FROM N.A.
MEDLINE=98038990; PubMed=9373153;
Thorpe K.L., Gorman P., Thomas C., Sheer D., Trowsdale J., Beck S.;
"Chromosomal localization, gene structure and transcription pattern of
the ORFX gene, a homologue of the MHC-linked RING3 gene.";
Gene 200:177-183 (1997).
-!- SUBCELLULAR LOCATION: Nuclear (Potential).
-!- TISSUE SPECIFICITY: Ubiquitous.
-!- SIMILARITY: Contains 2 bromodomains.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its

```



CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
DR EMBL; D26362; BAA05393.1; -  
DR EMBL; Z81330; -; NOT\_ANNOTATED\_CDS.  
DR HSP; Q92831; 1B91.  
DR Genew; HGNC:1104; BRD3.  
DR MY; 601541; -  
DR GO; GO:0005634; C:nucleus; NAS.  
DR InterPro; IPR001487; Bromodomain.  
DR Pfam; PF00439; bromodomain; 2.  
DR PRINTS; PR00503; BROMODOMAIN.  
DR SMART; SM00297; BROMO; 2.  
DR PROSITE; PS00633; BROMODOMAIN 1; 2.  
DR PROSITE; PS00633; BROMODOMAIN 2; 2.  
DR Bromodomain; Repeat; Nuclear protein.  
FT DOMAIN 56 115 BROMODOMAIN 1.  
FT DOMAIN 326 398 BROMODOMAIN 2.  
FT DOMAIN 487 555 LYS-RICH.  
FT DOMAIN 676 725 SER-RICH.  
FT CONFLICT 465 466 EL -> DV (IN REF. 2).  
SQ SEQUENCE 726 AA; 79541 MW; 64F526FC3C1033AA CRC64;

Query Match 85.3%; Score 110; DB 1; Length 726;  
Best Local Similarity 83.3%; Pred. No. 1.2e-09;  
Matches 20; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 FROPVDAVKLGDPYHKIKQPM 24  
DB 59 FYQFVDAIKNLDPYHKIKNPM 82

## RESULT 4

ID BRD4 HUMAN STANDARD; PRT; 1362 AA.  
AC OC0835; Q96PD3;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Bromodomain-containing protein 4 (HUNK1 protein).  
GN BRD4 OR HUNK1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21590020; PubMed=11733348;  
RA French C.A., Miyoshi I., Aster J.C., Kubonishi I., Kroll T.G.,  
RA Dai C.H., Vargas S.O., Perez-Atayde A.R., Fletcher J.A.;  
RT "BRD4 bromodomain gene rearrangement in aggressive carcinoma with  
RT translocation t(15;19).";  
RL Am. J. Pathol. 159:1987-1992 (2001).  
RN [2]  
RP SEQUENCE OF 1-722 FROM N.A.  
RC TISSUE=Placenta;  
RA Weber B.;  
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.  
CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).  
CC -1- SIMILARITY: Contains 2 bromodomains.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
DR EMBL; AF386649; AAL26987.1; -  
DR EMBL; Y12059; CAA72780.1; -

DR HSP; Q92831; 1B91.  
DR Genew; HGNC:13575; BRD4.  
DR InterPro; IPR001487; Bromodomain.  
DR Pfam; PF00439; bromodomain; 2.  
DR PRINTS; PR00503; BROMODOMAIN.  
DR SMART; SM00297; BROMO; 2.  
DR PROSITE; PS00633; BROMODOMAIN 1; 1.  
DR PROSITE; PS00633; BROMODOMAIN 2; 2.  
DR Bromodomain; Repeat; Nuclear protein.  
FT DOMAIN 75 147 BROMODOMAIN 1.  
FT DOMAIN 368 440 BROMODOMAIN 2.  
FT DOMAIN 535 594 LYS-RICH.  
FT DOMAIN 692 717 SER-RICH.  
FT DOMAIN 703 714 POLY-SER.  
FT DOMAIN 738 743 POLY-HIS.  
FT DOMAIN 757 761 POLY-PRO.  
FT DOMAIN 764 770 POLY-PRO.  
FT DOMAIN 771 775 POLY-GLN.  
FT DOMAIN 776 783 POLY-PRO.  
FT DOMAIN 954 964 POLY-PRO.  
FT DOMAIN 974 986 POLY-PRO.  
FT DOMAIN 1011 1014 POLY-PRO.  
FT DOMAIN 1028 1033 POLY-PRO.  
FT DOMAIN 1283 1300 POLY-GLN.  
FT DOMAIN 1301 1308 POLY-ALA.  
FT DOMAIN 1335 1338 POLY-ARG.  
FT CONFLICT 720 721 EM -> GP (IN REF. 2).  
SQ SEQUENCE 1362 AA; 152219 MW; D52EFCF9960907 CRC64;

Query Match 82.9%; Score 107; DB 1; Length 1362;  
Best Local Similarity 83.3%; Pred. No. 7.3e-09;  
Matches 20; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 FROPVDAVKLGDPYHKIKQPM 24  
DB 83 FQFVDAVKNLDPYHKIKTPM 106

## RESULT 5

ID CBPI CAEEL STANDARD; PRT; 2056 AA.  
AC P34545;  
DT 01-FEB-1994 (Rel. 28, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Protein cbp-1.  
GN CBP-1 OR R10E11.1.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Peleoderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RX MEDLINE=94150718; PubMed=7906398;  
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M., Coulson A.,  
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Fraser A.,  
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,  
RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,  
RA Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,  
RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,  
RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkhen R.,  
RA Sims M., Smaldon N., Smith A., Smith M., Sonnenhammer E., Staden R.,  
RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,  
RA Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,  
RA Wolfdmnan P.;  
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
RT elegans.";  
RL Nature 368:32-38 (1994).  
RN [2]  
RP REVISIONS, AND ALTERNATIVE SPLICING.

RA Durbin R.;  
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.

```

CC -|- ALTERNATIVE PRODUCTS;
CC Event=Alternative splicing; Named isoforms=2;
CC Name=b;
CC IsoId=P34545-1; Sequence=Displayed;
CC Name=a;
CC IsoId=P34545-2; Sequence=VSP_000557;
CC Note=No experimental confirmation available;
CC -|- SIMILARITY: Contains 1 bromodomain.
CC -|- SIMILARITY: Contains 1 ZF-type zinc finger.
CC -|- SIMILARITY: Contains 2 TAZ-type zinc fingers.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z23095; CAA82353.2; -
CC EMBL; Z23095; CAD18875.1; -
CC PIR; G88564; G88564.
CC WormPep; R10E11.1a; CE28069.
CC WormPep; R10E11.1b; CE21117.
CC InterPro; IPR001487; Bromodomain.
CC InterPro; IPR003101; KIX.
CC InterPro; IPR00197; TAZ_finger.
CC InterPro; IPR001965; Znf_PHD.
CC Pfam; PF00439; bromodomain; 1.
CC Pfam; PF02172; KIX; 1.
CC Pfam; PF02135; zf-TAZ; 2.
CC Pfam; PF00569; ZF; 1.
CC PRINTS; PR00503; BROMODOMAIN.
CC SMART; SM00297; BROMO; 1.
CC SMART; SM00551; Znf_TAZ; 2.
CC SMART; SM00291; Znf_ZZ; 1.
CC PROSITE; PS00633; BROMODOMAIN_1; 1.
CC PROSITE; PS00014; BROMODOMAIN_2; 1.
CC PROSITE; PS01359; ZF_PHD_1; 1.
CC PROSITE; PS01359; ZF_PHD_2; 1.
CC PROSITE; PS01357; ZF_TAZ; 2.
CC PROSITE; PS01357; ZF_ZZ_1; 1.
CC PROSITE; PS01357; ZF_ZZ_2; 1.
CC Bromodomain; Metal-binding; Zinc; Zinc-finger; Alternative splicing;
KW Repeat.
FT ZN_FING 399 505 TAZ-TYPE 1.
FT DOMAIN 881 953 BROMODOMAIN.
FT ZN_FING 1493 1534 ZZ-TYPE.
FT ZN_FING 1550 1631 TAZ-TYPE 2.
FT DOMAIN 1687 2008 GLY/GLN-RICH.
FT VARSPLIC 467 478 SDPTQTKKGV -> F (in isoform a).
FT FTID=VSP_000557.
SQ SEQUENCE 2056 AA; 227179 MW; 949FF4608C634F01 CRC64;
Query Match 72.1%; Score 93; DB 1; Length 2056;
Best Local Similarity 70.8%; pred. No. 2e-06;
Matches 17; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
Qy 1 FRPVDKVLGLPDYHKIKQPM 24
||| ||| :|||: |||
Db 889 FRPVDKLNIPDYHEIKRPM 912
-----
RESULT 6
P300 HUMAN
ID_P300 HUMAN STANDARD; PRT; 2414 AA.
AC Q09472;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE EIA-associated protein p300 (EC 2.3.1.48).
GN EP300 OR P300.
OS Homo sapiens (Human).

```

```

OC CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OX CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1] NCBI_taxID=9606;
RP SEQUENCE FROM N.A.
RX MEDLINE=95011587; PubMed=7523245;
RA Eckner R., Ewen M.E., Newsome D., Gerdes M., Decaprio J.A.,
RT Lawrence J.B., Livingston D.M.;
RT "Molecular cloning and functional analysis of the adenovirus E1A-
RT associated 300-kD protein (p300) reveals a protein with properties of
RT a transcriptional adaptor."
RL Genes Dev. 8:869-884(1994).
[2]
RP ENZYMOLOGICAL ACTIVITY.
RX MEDLINE=97100994; PubMed=8945521;
RA Ogryzko V.V., Schiltz R.L., Russanova V., Howard B.H., Nakatani Y.;
RT "The transcriptional coactivators p300 and CBP are histone
RT acetyltransferases."
RL Cell 87:953-959(1996).
[3]
RP INTERACTION WITH PCAF.
RX MEDLINE=96300317; PubMed=8684459;
RA Yang X.-J., Ogryzko V.V., Nishikawa J.-I., Howard B.H., Nakatani Y.;
RT "A p300/CBP-associated factor that competes with the adenoviral
RT oncoprotein E1A."
RL Nature 382:319-324(1996).
[4]
RP INTERACTION WITH HIF1A AND CREBBP.
RX MEDLINE=97075102; PubMed=8917528;
RA Arany Z., Huang L.E., Eckner R., Bhattacharya S., Jiang C.,
RT Goldberg M.A., Bunn H.F., Livingston D.M.;
RT "An essential role for p300/CBP in the cellular response to hypoxia."
RL Proc. Natl. Acad. Sci. U.S.A. 93:12969-12973(1996).
[5]
RP INTERACTIONS WITH NR3C1.
RX MEDLINE=98250578; PubMed=9590696;
RA Fryer C.J., Archer T.K.;
RT "Chromatin remodelling by the glucocorticoid receptor requires the
RT BRG1 complex."
RL Nature 393:88-91(1998).
[6]
RP INTERACTION WITH NCOA6.
RX MEDLINE=20283976; PubMed=10823961;
RA Ko L., Cardona G.R., Chin W.W.;
RT "Thyroid hormone receptor-binding protein, an LXXLL motif-containing
RT protein, functions as a general coactivator."
RL Proc. Natl. Acad. Sci. U.S.A. 97:6212-6217(2000).
[7]
RP INTERACTION WITH DTX1.
RX MEDLINE=21576166; PubMed=11564735;
RA Yamamoto N., Yamamoto S.-I., Inagaki F., Kawauchi M., Fukamizu A.,
RT Kishi N., Matsuno K., Nakamura K., Weinmaster G., Okano H.,
RA Nakafuku M.;
RT "Role of Deltex-1 as a transcriptional regulator downstream of the
RT Notch receptor."
RL J. Biol. Chem. 276:45031-45040(2001).
[8]
RP PHOSPHORYLATION, AND INTERACTION WITH TCF7L2 AND LEF1.
RX MEDLINE=22450614; PubMed=12446687;
RA Hecht A., Stemmer M.P.;
RT "Identification of a promoter-specific transcriptional activation
RT domain at the C-terminus of the Wnt-effector protein TCF4."
RL J. Biol. Chem. 278:3776-3785(2003).
[9]
RP VARIANTS BREAST CANCER PRO-827 AND GLY-1013, VARIANT PANCREATIC CANCER
RP TYR-1650, AND VARIANT COLORECTAL CANCER GLN-2221.
RX MEDLINE=20164332; PubMed=10700188;
RA Gayther S.A., Batley S.J., Linger L., Bannister A., Thorpe K.,
RT Chin S.-F., Daigo Y., Russell P., Wilson A., Sower H.M.,
RA Delhanty J.D.A., Ponder B.A.J., Kouzarides T., Caldas C.;
RT "Mutations truncating the EP300 acetylase in human cancers."
RL Nat. Genet. 24:300-303(2000).
[10]
RN

```

```

FT DOMAIN 2190 2195 POLY-GLN.
FT VARIANT 827 827 L -> P (in breast cancer).
FT FT /FTID=VAR 014428.
FT VARIANT 1013 1013 E -> G (in breast cancer).
FT FT /FTID=VAR 014429.
FT VARIANT 1650 1650 S -> Y (in pancreatic cancer).
FT FT /FTID=VAR 014430.
FT VARIANT 2221 2221 P -> Q (in colorectal cancer).
FT FT /FTID=VAR 014431.
FT SQ SEQUENCE 2414 AA; 264143 MW; 6BFF909EB4B9D693 CRC64;

Query Match 68.2%; Score 88; DB 1; Length 2414;
Best Local Similarity 66.7%; Pred. No. 1.5e-05;
Matches 16; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 FROPVDVAVKLGLPDYKIKIQPMND 24
||||| ||||| |||||
DB 1075 FROPVDPQLLGIPDYFDIVKSPMD 1098

RESULT 7
CBP_MOUSE
ID CBP_MOUSE STANDARD; PRT; 2441 AA.
AC P45481;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE C3EB-binding protein (EC 2.3.1.48).
GN C3EBBP OR CBP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=94019866; PubMed=8413673;
RA Carliva J.C., Kwok R.P.S., Lamb N., Hagiwara M., Montminy M.R.,
RA Goodnan R.H.;
RT "Phosphorylated CREB binds specifically to the nuclear protein CBP.";
RN Nature 365:855-859(1993).
RL [2]
RP INTERACTION WITH NCOA3.
RX MEDLINE=97336097; PubMed=9192892;
RA Torchia J., Rose D.W., Inostroza J., Kamei Y., Westin S., Glass C.K.,
RA Rosenfeld M.G.;
RT "The transcriptional co-activator p/Cip binds CBP and mediates
RT nuclear-receptor function.";
RN Nature 387:677-684(1997).
RL [2]
CC -1- FUNCTION: Acetyltransferase enzyme. Acetylates histones, giving a
CC specific tag for transcriptional activation. Also acetylates non-
CC histone proteins, like NCOA3 coactivator. Mediates CAMP-gene
CC regulation by binding specifically to phosphorylated CREB protein.
CC CBP, as coactivator, augments the activity of phosphorylated CREB
CC to activate transcription of CAMP-responsive genes (By
CC similarity).
CC -1- CATALYTIC ACTIVITY: Acetyl-CoA + histone = CoA + acetyl-histone.
CC -1- SUBUNIT: Interacts with SMAD1, SMAD2, SMAD3, PCAF and PML. The
CC TAZ-type 1 domain interacts with HIF1A. Found in a complex
CC containing NCOA2; NCOA3; IKKα; IKKβ and IKKγ. Interacts with
CC NCOA6 coactivator. Probably part of a complex with HIF1α and
CC EP300 (By similarity).
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: Contains 1 bromodomain.
CC -1- SIMILARITY: Contains 1 ZZ-type zinc finger.
CC -1- SIMILARITY: Contains 2 TAZ-type zinc fingers.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

or send an email to license@isb-sib.ch).

```

CC EMBL; S66385; AAB28651.1; -
CC PIR; S39161; S39161
DR PDB; 1F81; 18-OCT-00.
DR PDB; 1JUS; 03-OCT-01.
DR PDB; 1KBH; 06-FEB-02.
DR PDB; 1KDX; 25-NOV-98.
DR PDB; 1L8C; 24-APR-02.
DR TRANSFAC; T01318; -.
DR MGD; NGI.109280; Crebbp.
DR GO; GO:000123; C:histone acetyltransferase complex; IDA.
DR GO; GO:0005634; C:nucleus; IDA.
DR GO; GO:0004402; F:histone acetyltransferase activity; IDA.
DR GO; GO:0005515; F:protein binding; IPI.
DR GO; GO:0003713; F:transcription co-activator activity; IDA.
DR GO; GO:0045449; P:regulation of transcription; IDA.
DR InterPro; IPR001487; Bromodomain.
DR InterPro; IPR003101; KIX.
DR InterPro; IPR000197; TAZ_finger.
DR InterPro; IPR000433; ZnF_ZZ.
DR Pfam; PF00439; bromodomain; 1.
DR Pfam; PF02172; KIX; 1.
DR Pfam; PF02135; zf-TAZ; 2.
DR Pfam; PF00589; ZZ; 1.
DR PRINTS; PR00503; BROMODOMAIN.
DR SMART; SM00297; BROMO; 1.
DR SMART; SM00551; ZnF_TAZ; 2.
DR SMART; SM00291; ZnF_ZZ; 1.
DR PROSITE; PS00633; BROMODOMAIN_1; 1.
DR PROSITE; PS00014; BROMODOMAIN_2; 1.
DR PROSITE; PS00134; ZF_TAZ; 2.
DR PROSITE; PS01357; ZF_ZZ_1; 1.
DR PROSITE; PS0135; ZF_ZZ_2; 1.
DR PROSITE; PS00135; ZF_ZZ_2; 1.
KW Transferase; Transcription regulation; Nuclear protein; Activator;
KW Bromodomain; Zinc-finger; Repeat; 3D-structure.
FT ZN FING 346 432 TAZ-TYPE 1.
FT DOMAIN 1104 1176 BROMODOMAIN.
FT ZN FING 1702 1745 ZZ-TYPE.
FT ZN FING 1766 1847 TAZ-TYPE 2.
FT DOMAIN 1062 1065 POLY-GLU.
FT DOMAIN 1556 1563 POLY-GLU.
FT DOMAIN 1944 1949 POLY-PRO.
FT DOMAIN 1968 1971 POLY-GLN.
FT DOMAIN 2082 2086 POLY-GLN.
FT DOMAIN 2200 2216 POLY-GLN.
FT DOMAIN 2296 2299 POLY-GLN.
SQ SEQUENCE 2441 AA; 265474 MW; 0ABBO28C3112F419 CRC64;

Query Match 68.2%; Score 88; DB 1; Length 2441;
Best Local Similarity 66.7%; Pred. No. 1.5e-05;
Matches 16; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 1 PRQPVDAVKLGDPYHKKIKQPMW 24
      ||||| :||| :||| :|||
Db 1112 FRQVDPQLGIPDYFDVKNPMD 1135

RESULT 8
CBP_HUMAN
ID _CBP_HUMAN STANDARD; PRT; 2442 AA.
AC Q92793; Q00147; Q16376;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE CREB-binding protein (EC 2.3.1.48).
GN CREBBP OR CBP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97385172; PubMed=9238046;
RA Sobulo O.M., Borrow J., Tomek R., Reshimi S., Harden A.,
RA Schlegelberger B., Housman D., Doggett N.A., Rowley J.D.,
RA Zeleznik-Le N.J.;
RT "MLL is fused to CBP, a histone acetyltransferase, in therapy-related
RT acute myeloid leukemia with a t(11;16)(q23;p13.3).";
RL Proc. Natl. Acad. Sci. U.S.A. 94:8732-8737(1997).
RN [2]
RN SEQUENCE FROM N.A.
RX MEDLINE=97321049; PubMed=9177780;
RA Giles R.H., Petrij F., Dauwerse H.G., den Hollander A.I.,
RA Lushnikova T., van Ommen G.J.B., Goodman R.H., Deaven L.L.,
RA Doggett N.A., Peters D.J.M., Breuning M.H.;
RT "Construction of a 1.2-Mb contig surrounding, and molecular analysis
RT of, the human CREB-binding protein (CBP/CREBBP) gene on chromosome
RT 16p13.3.";
RL Genomics 42:96-144(1997).
RN [3]
RN SEQUENCE OF 1-405 FROM N.A.
RX MEDLINE=96378968; PubMed=8782817;
RA Borrow J., Stanton V.P., Andresen J.M., Becher R., Behm F.G.,
RA Chaganti R.S.K., Civin C.I., Distche C., Dube I., Frischau A.M.,
RA Horsman D., Mitelman P., Volinia S., Watmore A.E., Housman D.E.;
RT "The translocation t(8;16)(p11;p13) of acute myeloid leukaemia fuses
RT a putative acetyltransferase to the CREB-binding protein.";
RL Nat. Genet. 14:33-41(1996).
RN [4]
RN INTERACTION WITH PCAF.
RX MEDLINE=96300317; PubMed=8684459;
RA Yang X.-J., Ogryzko V.V., Nishikawa J.-I., Howard B.H., Nakatani Y.;
RT "A p300/CBP-associated factor that competes with the adenoviral
RT oncoprotein E1A.";
RL Nature 382:319-324(1996).
RN [5]
RN INTERACTION WITH HIF1A AND EP300.
RX MEDLINE=97075102; PubMed=8917528;
RA Arany Z., Huang L.E., Eckner R., Bhattacharya S., Jiang C.,
RA Goldberg M.A., Bunn H.F., Livingston D.M.;
RT "An essential role for p300/CBP in the cellular response to hypoxia.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:12969-12973(1996).
RN [6]
RN INTERACTION WITH PML.
RX MEDLINE=99178939; PubMed=10077561;
RA Doucas V., Tini M., Egan D.A., Evans R.M.;
RT "Modulation of CREB binding protein function by the promyelocytic
RT (PML) oncoprotein suggests a role for nuclear bodies in hormone
RT signaling.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:2627-2632(1999).
RN [7]
RN ACETYLATION OF NCOA3.
RX MEDLINE=99418638; PubMed=10490106;
RA Chen H., Lin R.J., Xie W., Wilpitz D., Evans R.M.;
RT "Regulation of hormone-induced histone hyperacetylation and gene
RT activation via acetylation of an acetylase.";
RL Cell 98:675-686(1999).
RN [8]
RN INTERACTION WITH NCOA6.
RX MEDLINE=20325329; PubMed=10866662;
RA Mahajan M.A., Samuels H.H.;
RT "A new family of nuclear receptor coregulators that integrates nuclear
RT receptor signaling through CBP.";
RL Mol. Cell. Biol. 20:5048-5063(2000).
RN [9]
RN IDENTIFICATION IN A COMPLEX WITH NCOA2; NCOA3; IKK $\alpha$ ; IKK $\beta$  AND IKK $\gamma$ .
RX MEDLINE=21968797; PubMed=11971985;
RA Wu R.C., Qin J., Hashimoto Y., Wong J., Xu J., Tsai S.Y., Tsai M.J.,
RA O'Malley B.W.;
RT "Regulation of SRC-3 (pCIP/ACTR/AIB-1/RAC-3/TRAM-1) coactivator
RT activity by I $\kappa$ B kinase.";
RL Mol. Cell. Biol. 22:3549-3561(2002).
RN [10]
RN STRUCTURE BY NMR OF 345-439 IN COMPLEX WITH 776-826 OF HIF1A.
RX MEDLINE=21957241; PubMed=11959977;

```

RA Dames S.A., Martinez-Yamout M., De Guzman R.N., Dyson H.J.,  
RA Wright P.E.;  
RA "Structural basis for Hif-1 alpha /CBP recognition in the cellular  
RA hypoxic response.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:5271-5276(2002).  
RF [11]  
RN VARIANT RSTS PRO-1378.  
RX MEDLINE=21231140; PubMed=11331617;  
RA Murata T., Kurokawa R., Krones A., Tatsumi K., Ishii M., Taki T.,  
RA Masuno M., Ohashi H., Yanagisawa M., Rosenfeld M.G., Glass C.K.,  
RA Hayashi Y.;  
RA "Defect of histone acetyltransferase activity of the nuclear  
RT transcriptional coactivator CBP in Rubinstein-Taybi syndrome.";  
RL Hum. Mol. Genet. 10:1071-1076(2001).  
CC -|- FUNCTION: Acetyltransferase enzyme. Acetylates histones, giving a  
CC specific tag for transcriptional activation. Also acetylates non-  
CC histone proteins, like NCOA3 coactivator. Mediates cAMP-gene  
CC regulation by binding specifically to phosphorylated CREB protein.  
CC CBP, as coactivator, augments the activity of phosphorylated CREB  
CC to activate transcription of cAMP-responsive genes.  
CC -|- CATALYTIC ACTIVITY: Acetyl-CoA + histone = CoA + acetyl-histone.  
CC -|- SUBUNIT: Interacts with SMAD1, SMAD2, SMAD3, PCAF and PML. The  
CC TAZ-type 1 domain interacts with HIF1A. Found in a complex  
CC containing NCOA2; NCOA3; IKKX; IKKB and IKKG. Interacts with  
CC NCOA6 coactivator. Probably part of a complex with Hif1a and  
CC BOP30.  
CC -|- SUBCELLULAR LOCATION: Nuclear.  
CC -|- DISEASE: Involved in acute myeloid leukemias through chromosomal  
CC translocations t(8;16)(p11;p13) involving MYST3 and CREBBP,  
CC and t(11;16)(q23;p13.3) involving MLL/HRX and CREBBP.  
CC -|- DISEASE: Defects in CREBBP are the cause of Rubinstein-Taybi  
CC syndrome (RSTS) [MIM:180849]. RSTS is an autosomal dominant  
CC disorder characterized by craniofacial abnormalities, broad  
CC thumbs, broad big toes, mental retardation and a propensity for  
CC development of malignancies.  
CC -|- SIMILARITY: Contains 1 bromodomain.  
CC -|- SIMILARITY: Contains 2 TAZ-type zinc finger.  
CC -|- SIMILARITY: Contains 2 TAZ-type zinc fingers.  
CC -|- DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.;  
CC WWW="http://www.infobiogen.fr/services/chronocancer/Genes/CBPID42.html".  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC -----  
CC EMBL; U47741; AAC51770.1; --  
CC EMBL; U85962; AAC51331.1; --  
CC EMBL; U89354; AAC51339.1; --  
CC EMBL; U89355; AAC51340.1; --  
CC DB; 1L8C; 24-APR-02.  
CC DB; 1JSP; 17-AUG-02.  
CC TRANSPAC; T02214; --  
CC Genew; HGNC:2348; CREBBP.  
CC MIM; 600140; --  
CC MIM; 180849; --  
CC GO; GO:0005737; Cytoplasm; TAS.  
CC GO; GO:0005634; C:nucleus; TAS.  
CC GO; GO:0004871; F:signal transducer activity; TAS.  
CC GO; GO:0003713; F:transcription co-activator activity; TAS.  
CC GO; GO:0003700; F:transcription factor activity; TAS.  
CC GO; GO:0006461; P:protein complex assembly; TAS.  
CC GO; GO:0006355; P:protein complex transcription, DNA-dependent; TAS.  
CC GO; GO:0007165; P:signal transduction; TAS.  
CC InterPro; IPR001487; Bromodomain.  
CC InterPro; IPR003101; KIX.  
CC InterPro; IPR000197; TAZ finger.  
CC InterPro; IPR000433; Znf\_ZZ.  
CC Pfam; PF00439; bromodomain; 1.  
CC Pfam; PF02172; KIX; 1.

DR Pfam; PF02135; zf-TAZ; 2.  
DR Pfam; PF00569; ZZ; 1.  
DR PRINIS; PRO0503; BROMODOMAIN.  
DR SMART; SM00297; BROMO; 1.  
DR SMART; SM00551; Znf\_TAZ; 2.  
DR SMART; SM00291; Znf\_ZZ; 1.  
DR PROSITE; PS00633; BROMODOMAIN\_1; 1.  
DR PROSITE; PS00134; BROMODOMAIN\_2; 1.  
DR PROSITE; PS00134; ZF\_TAZ; 2.  
DR PROSITE; PS00357; ZF\_ZZ\_1; 1.  
DR PROSITE; PS00135; ZF\_ZZ\_2; 1.  
KW Transferase; Transcription regulation; Nuclear protein; Activator;  
KW Bromodomain; Chromosomal translocation; Zinc-finger; Repeat;  
KW Disease mutation; 3D-structure.  
FT ZN FING 1701 1744 ZZ-TYPE.  
FT ZN FING 347 433 TAZ-TYPE 1.  
FT DOMAIN 452 683 CREB-BINDING.  
FT DOMAIN 1103 1175 BROMODOMAIN.  
FT DOMAIN 1061 1064 POLY-GLU.  
FT DOMAIN 1199 1487 CYS/HIS-RICH.  
FT DOMAIN 1555 1562 POLY-GLU.  
FT ZN FING 1765 1846 TAZ-TYPE 2.  
FT DOMAIN 1943 1948 POLY-PRO.  
FT DOMAIN 1967 1970 POLY-GLN.  
FT DOMAIN 2081 2085 POLY-GLN.  
FT DOMAIN 2199 2216 POLY-GLN.  
FT DOMAIN 2245 2248 POLY-GLN.  
FT DOMAIN 2297 2300 POLY-GLN.  
FT VARIANT 1378 1378 R -> P (in RSTS; abolishes  
FT acetyltransferase activity and the  
FT ability of transactivate CREB).  
FT /FTID=VAR\_015578.  
FT CONFLICT 1511 1513 FAE -> NSG (IN REF. 2).  
FT CONFLICT 1724 1725 ED -> VV (IN REF. 2).  
FT CONFLICT 1770 1770 V -> L (IN REF. 2).  
FT CONFLICT 1789 1789 N -> F (IN REF. 2).  
FT CONFLICT 1812 1812 T -> P (IN REF. 2).  
SQ SEQUENCE 2442 AA; 285336 MW; 42D084619475F3D2 CRC64;  
  
Query Match 68.2%; Score 88; DB 1; Length 2442;  
Best Local Similarity 66.7%; Pred. No. 1.5e-05;  
Matches 16; Conservative 2; Mismatches 6; Indels 0; Gaps 0;  
  
QY 1 PRQVDAVKLGPDYHKIKOPMD 24  
DB 1111 PRQVDPOLGIPDYFDIVKNPMD 1134  
||||| :||| :|||  
1 PRQVDAVKLGPDYHKIKOPMD 24  
1111 PRQVDPOLGIPDYFDIVKNPMD 1134  
  
RESULT 9  
YK82 SCHPO  
ID YK82 SCHPO STANDARD; PRT; 727 AA.  
AC Q9HGF4; P78808;  
DT 10-OCT-2003 (Rel. 42, Created)  
DT 10-OCT-2003 (Rel. 42, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Hypothetical bromodomain protein C631.02.  
GN SPAC631.02.  
OS Schizosaccharomyces pombe (Fission yeast).  
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
OC Schizosaccharomycetes.  
ON NCBI\_TaxID=4896;  
RX [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=972;  
RX MEDLINE=21848401; PubMed=11859360;  
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,  
RA Gentles S., Goble A., Hamlin N., Harris D., Hidaigo J., Hodgson G.,  
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,

```

RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharps S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritzc C., Holzer E., Moestl D., Hilbert H.,
RA Borzpm K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler E., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Delaure V., Mottier S.,
RA Galibert F., Aves S.J., Kiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Spakevski G.V., Ussery D., Barrell B.G., Nurse P.,
RT "The genome sequence of Schizosaccharomyces pombe."
RL Nature 415:871-880 (2002).
RN [2]
RC STRAIN=PR745;
RC STRAIN=5288c / AB972;
RC MEDLINE=98162722; PubMed=9501991;
RX Yoshioka S., Kato K., Nakai K., Okayama H., Nojima H.;
RA "Identification of open reading frames in Schizosaccharomyces pombe
RT cDNAs."
RL DNA Res. 4:363-369 (1997).
CC -!- SIMILARITY: Contains 2 bromodomains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AL391713; CAC05484.1; -.
DR EMBL; D89157; BAA13819.1; -.
DR PIR; T42517; T42517.
DR GeneDB SPombe; SPAC631.02; -.
DR InterPro; IPR001487; Bromodomain.
DR Pfam; PF00439; bromodomain; 2.
DR PRINTS; PR00503; BROMODOMAIN.
DR SMART; SM00297; BROMO; 2.
DR PROSITE; PS00633; BROMODOMAIN 1; 1.
DR PROSITE; PS00633; BROMODOMAIN 2; 2.
DR Hypothetical protein; Bromodomain; Repeat.
FT DOMAIN 246 318 BROMODOMAIN 1.
FT DOMAIN 408 480 BROMODOMAIN 2.
SQ SEQUENCE 727 AA; 81031 MW; 5CFA73844CB6EF05 CRC64;

Query Match 61.2%; Score 79; DB 1; Length 727;
Best Local Similarity 62.5%; Pred. No. 0.00011;
Matches 15; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 FROPVDVAVKGLPDYHKIKQPMQ 24
DQ 254 FRAPVDVAVKQNPDPYFIKNPFD 277

RESULT 10
BD1 YEAST
ID BD1 YEAST STANDARD; PRT; 686 AA.
AC P35817; Q06048;
DT 01-JUN-1994 (Rel. 29, Created)
DT 15-JUN-1999 (Rel. 38, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE BD1 protein.
GN BD1 OR YLR399C OR L8084.18.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.

```

```

OX NCBI_TaxID=4932;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=S288c;
RX MEDLINE=95116323; PubMed=7816623;
RA Lygerou Z., Conesa C., Lesage P., Swanson R.N., Ruet A., Carlson M.,
RA Senterac A., Seraphin B.;
RT "The yeast BDF1 gene encodes a transcription factor involved in the
RT expression of a broad class of genes including snRNAs."
RL Nucleic Acids Res. 22:5332-5340 (1994).
RN [2]
RN SEQUENCE FROM N.A.
RA Roeder G.S., Chua P.R.;
RA Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RN SEQUENCE FROM N.A.
RC STRAIN=S288c / AB972;
RX MEDLINE=97313267; PubMed=9169871;
RA Johnston M., Hillier L., Riles L., Albermann K., Andre B., Ansorge W.,
RA Benes V., Brueckner M., Delius H., Dubois E., Duesterhoeft A.,
RA Entian K.-D., Floeth M., Goffeau A., Hebling U., Heumann K.,
RA Heuss-Neitzel D., Hilbert H., Hilger F., Kleine K., Koetter P.,
RA Louis E.J., Messenguy F., Mewes H.-W., Miosga T., Moestl D.,
RA Mueller-Auer S., Nentwich U., Obermaier B., Piravandi E., Pohl T.M.,
RA Portetelle D., Purnelle B., Rechmann S., Rieger M., Rinke M., Rose M.,
RA Scharfe M., Scherens B., Scholler P., Schwager C., Schwarz S.,
RA Underwood A.P., Urrestazu L.A., Vandenbol M., Verhasselt P.,
RA Vierendeels F., Voet M., Volckaert G., Voss H., Wambutt R., Wedler E.,
RA Wedler H., Zimmermann F.K., Zollner A., Hani J., Hoheisel J.D.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome XII."
RL Nature 387:87-90 (1997).
RN [4]
RN SEQUENCE OF 471-686 FROM N.A.
RX MEDLINE=93309467; PubMed=8321235;
RA Widner W.R., Wickner R.B.;
RT "Evidence that the SKI antiviral system of Saccharomyces cerevisiae
RT acts by blocking expression of viral mRNA."
RL Mol. Cell. Biol. 13:4331-4341 (1993).
CC -!- FUNCTION: Transcription factor involved in the expression of a
CC -!- broad class of genes including snRNAs. Required for sporulation.
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -!- SIMILARITY: Contains 2 bromodomains.
CC -!- SIMILARITY: Contains 1 ET domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z18944; CAA79377.1; -.
DR EMBL; U08116; AAB89115.1; -.
DR EMBL; U19729; AAB82357.1; -.
DR EMBL; L13469; AAA35048.1; -.
DR PIR; S55955; S55955.
DR GenOnline; 142462; -.
DR TRANSFAC; T03204; -.
DR SGD; S0004391; BDF1.
DR GO; GO:0005634; C:nucleus; IDA.
DR GO; GO:0030528; F:transcription regulator activity; IMP.
DR GO; GO:0007151; P:sporulation (sensu Saccharomyces); IMP.
DR InterPro; IPR001487; Bromodomain.
DR Pfam; PF00439; bromodomain; 2.
DR PRINTS; PR00503; BROMODOMAIN.
DR SMART; SM00297; BROMO; 2.
DR PROSITE; PS00633; BROMODOMAIN 1; 2.
DR PROSITE; PS00633; BROMODOMAIN 2; 2.
KW Transcription regulation; Nuclear protein; Bromodomain; Repeat;
KW Sporulation.
FT DOMAIN 165 237 BROMODOMAIN 1.
FT DOMAIN 332 404 BROMODOMAIN 2.

```

```

FT DOMAIN 522 686 ET DOMAIN.
FT CONFLICT 8 8 Q -> LC (IN REF. 1).
FT CONFLICT 93 94 GA -> R (IN REF. 2).
FT CONFLICT 94 94 A -> P (IN REF. 2).
FT CONFLICT 282 282 A -> P (IN REF. 1).
FT CONFLICT 385 385 D -> E (IN REF. 1).
FT CONFLICT 493 493 A -> R (IN REF. 4).
SQ SEQUENCE 686 AA; 79978 MW; 8CDB52FA1F91D0DA CRC64;

Query Match 58.9%; Score 76; DB 1; Length 686;
Best Local Similarity 62.5%; Pred. No. 0.00033;
Matches 15; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 FRQVDVAVKLGIDYVHKIKQPMND 24
Db 173 FLOPVDVAVKLGIDYVHKIKQPMND 196

RESULT 11
ID_TF1A_HUMAN STANDARD; PRT; 1050 AA.
AC Q15164; O95854;
DT 15-JUL-1999 (Rel. 38, Created)
DI 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Transcription intermediary factor 1-alpha (TIF1-alpha) (Tripartite
DE motif protein 24).
GN TIF1 OR TRIM24 OR TIF1A.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM SHORT).
RC TISSUE=Breast cancer;
RX MEDLINE=97277352; PubMed=9115274;
RA Thenot S., Henriquet C., Rochefort H., Cavailles V.;
RA Venturini L., You J., Stadler M., Gallien R., Lallemand V.,
RA Koken M.H.M., Mattei M.-G., Ganser A., Chambon P., Losson R.,
RA De The H.;
RT "TIF1gamma, a novel member of the transcriptional intermediary factor
RT 1 family.";
RL Oncogene 18:1209-1217(1999).
RN [3]
RP SEQUENCE OF 477-510 (ISOFORM LONG).
RC TISSUE=Breast cancer;
RA Cavailles V.;
RL Submitted (JAN-1999) to Swiss-Prot.
RN [4]
RP INTERACTION WITH NR3C2.
RX MEDLINE=20388588; PubMed=10935545;
RA Hellal-Levy C., Fagart J., Souque A., Wurtz J.-M., Moras D.,
RA Rafestin-Obelin M.-E.;
RT "Crucial role of the H11-H12 loop in stabilizing the active
RT conformation of the human mineralocorticoid receptor.";
RL Mol. Endocrinol. 14:1210-1221(2000).
CC CC
CC -1- FUNCTION: Interacts selectively in vitro with the AF2-activating
CC domain of the estrogen receptors. Association with DNA-bound
CC estrogen receptors requires the presence of estradiol.
CC -1- SUBUNIT: Interacts with CBX1 and CBX3 (By similarity). Interacts
CC with NR3C2.
CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -1- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=2;
CC Name=Long;
CC IsoId=O15164-1; Sequence=Displayed;
CC Name=Short;

```

```

CC CC IsoId=O15164-2; Sequence=VSP_005772;
CC -1- SIMILARITY: Contains 1 RING-type zinc finger.
CC -1- SIMILARITY: Contains 2 B box-type zinc fingers.
CC -1- SIMILARITY: Contains 1 bromodomain.
CC -1- SIMILARITY: Contains 1 PHD-type zinc finger.
CC CC
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/ebis-sib.ch/
CC or send an email to license@ebi.ac.uk).
CC CC
CC EMBL; AF009353; AAB63585.1; -.
CC EMBL; AF119042; AAD17258.1; -.
CC HSSP; P29590; 1BOR.
CC TRANSFAC; T04945; -.
CC Genew; HGNC:11812; TIF1.
CC MIM; 603406; -.
CC GO; GO:0005634; C:nucleus; TAS.
CC GO; GO:0005102; F:receptor binding; TAS.
CC GO; GO:0003713; P:transcription co-activator activity; TAS.
CC GO; GO:0006366; P:transcription from Pol II promoter; TAS.
CC InterPro; IPR003649; BBox_C.
CC InterPro; IPR001487; Bromodomain.
CC InterPro; IPR000315; Znf_Bbox.
CC InterPro; IPR001965; Znf_PHD.
CC InterPro; IPR001841; Znf_ring.
CC Pfam; PF00439; bromodomain; 1.
CC Pfam; PF00628; PHD; 1.
CC Pfam; PF00643; zf-B_box; 2.
CC Pfam; PF00097; zf-C3HC4; 1.
CC PRINTS; PR01406; BBOXZNFINGER.
CC PRINTS; PR00503; BROMODOMAIN.
CC SMART; SM00502; BBC; 1.
CC SMART; SM00336; BBOX; 2.
CC SMART; SM00297; BROMO; 1.
CC SMART; SM00249; PHD; 1.
CC SMART; SM00184; RING; 1.
CC PROSITE; PS00633; BROMODOMAIN_1; FALSE_NEG.
CC PROSITE; PS00014; BROMODOMAIN_2; 1.
CC PROSITE; PS00119; ZF_BBOX; 2.
CC PROSITE; PS01359; ZF_PHD_1; 1.
CC PROSITE; PS00016; ZF_PHD_2; 1.
CC PROSITE; PS00518; ZF_RING_1; 1.
CC PROSITE; PS00089; ZF_RING_2; 1.
CC Transcription regulation; Repressor; DNA-binding; Bromodomain;
CC Zinc-finger; Alternative splicing; Nuclear protein; Coiled coil;
CC Repeat.
KW Repeat.
KW DOMAIN 9 15 POLY-ALA.
KW FT ZN FING 56 82 RING-TYPE.
KW FT ZN FING 158 211 B BOX-TYPE 1.
KW FT ZN FING 218 259 B BOX-TYPE 2.
KW FT DOMAIN 289 359 COILED COIL (POTENTIAL).
KW FT DOMAIN 344 347 POLY-GLN.
KW FT DOMAIN 754 779 NUCLEAR RECEPTOR BINDING SITE (NRBS).
KW FT ZN FING 826 873 PHD-TYPE.
KW FT DOMAIN 891 907 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
KW FT DOMAIN 932 987 Bromodomain.
KW FT VARSPLIC 477 510 Missing (in isoform Short).
KW FT /FTID=VSP_005772.
KW FT AASAAAS -> RLGCAP (IN REF. 1).
KW FT SAAPS -> RGG (IN REF. 1).
KW FT GSPVSG -> ARRSA (IN REF. 1).
KW FT A -> T (IN REF. 1).
KW FT D -> N (IN REF. 1).
KW FT M -> I (IN REF. 1).
KW FT A -> R (IN REF. 1).
SQ SEQUENCE 1050 AA; 116831 MW; D341E8022AAC67E CRC64;

Query Match 53.5%; Score 69; DB 1; Length 1050;
Best Local Similarity 68.4%; Pred. No. 0.0068;

```





CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

CC EMBL; AJ238214; CAC37033.2; --

CC EMBL; AL163279; CAB90452.1; --

CC EMBL; AJ222636; BAA10896.1; --

CC EMBL; AX002177; BAA9123.1; --

CC HSSP; Q28831; 1B91.

CC Genew; HGNC:12760; WDR9.

CC InterPro; IPR001487; Bromodomain.

CC Pfam; PF00439; bromodomain; 2.

CC Pfam; PF00400; WD40; 8.

CC PRINTS; PRO0503; BROMODOMAIN.

CC ProDom; PD000018; WD40; 1.

CC SMART; SM00297; BROMO; 2.

CC SMART; SM00320; WD40; 8.

CC PROSITE; PS00578; WD\_REPEATS\_1; 2.

CC PROSITE; PS00682; WD\_REPEATS\_2; 6.

CC PROSITE; PS00294; WD\_REPEATS\_REGION; 1.

CC PROSITE; PS00633; BROMODOMAIN\_1; FALSE\_NEG.

CC PROSITE; PS00114; BROMODOMAIN\_2; 2.

KW Bromodomain; Repeat; WD repeat.

FT REPEAT 184 223 WD 1.

FT REPEAT 226 255 WD 2.

FT REPEAT 286 311 WD 3.

FT REPEAT 322 365 WD 4.

FT REPEAT 366 405 WD 5.

FT REPEAT 424 463 WD 6.

FT REPEAT 466 506 WD 7.

FT REPEAT 514 553 WD 8.

FT DOMAIN 1177 1247 BROMODOMAIN 1.

FT DOMAIN 1330 1400 BROMODOMAIN 2.

FT DOMAIN 661 668 POLY-GLN.

FT DOMAIN 821 824 POLY-SER.

FT DOMAIN 1539 1548 POLY-SER.

FT DOMAIN 1687 1692 POLY-GLU.

FT CONFLICT 1560 1562 SSS -> GTR (IN REF. 3).

FT CONFLICT 1659 1659 P -> L (IN REF. 2).

FT CONFLICT 1752 1779 SHDSDHACNRVATGAPSTSV -> VMQIMHVQNCWPINVCQ (IN REF. 3).

FT CONFLICT 2134 2134 E -> G (IN REF. 4).

FT SEQUENCE 2259 AA; 257220 MW; 276D6ABE68C231F2 CRC64;

Query Match 52.7%; Score 68; DB 1; Length 2269;

Best Local Similarity 62.15%; Pred.No. 0.023;

Matches 15; Conservative 1; Mismatches 6; Indels 2; Gaps 1;

QY 1 FRQPVDAVKLGFLPDYHKIIKQPMD 24  
|||||: |||||

Db 1338 FRQPVDLVE--YPDYRDIIDIPMD 1359  
|||||: |||||

RESULT 14

TF1G HUMAN STANDARD; PRT; 1127 AA.

ID AC Q9UN9; Q9C017; Q9UJ79;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Transcription intermediary factor 1-gamma (TIF1-gamma) (RET-fused gene 7 protein) (Rif7 protein) (Tripartite motif protein 33).

GN TRIM33 OR TIF1G OR RFG7 OR KIAA1113.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

NCBI\_TaxID=9606;

RN [1]

SEQUENCE FROM N.A. (ISOFORM ALPHA).

RP MEDLINE=99144725; PubMed=10022127;

RX

RA Venturini L., You J., Stadler M., Gallien R., Lallemand V.,  
RA Koken M.H.M., Mattei M.-G., Ganter A., Chambon P., Losson R.,  
RA De The H.,  
RT "TIFgamma, a novel member of the transcriptional intermediary factor  
RT 1 family.";  
RL Oncogene 18:1209-1217(1999).  
RN [2]  
RN SEQUENCE FROM N.A. (ISOFORMS ALPHA AND BETA).  
RX MEDLINE=21231161; PubMed=11331580;  
RA Raymond A., Meroni G., Pantozzi A., Merla G., Cairo S., Luzzi L.,  
RA Risanello D., Zanaria E., Messali S., Cainarca S., Guffanti A.,  
RA Minucci S., Pelicci P.G., Ballabio A.,  
RT "The tripartite motif family identifies cell compartments.";  
RL EMBO J. 20:2140-2151(2001).  
RN [3]  
RN SEQUENCE FROM N.A. (ISOFORM ALPHA).  
RP TISSUE=Brain;  
RC MEDLINE=99397452; PubMed=10470851;  
RA Kikuno R., Nagase T., Ishikawa K.-I., Hirose M., Miyajima N.,  
RA Tanaka A., Kotani H., Nomura N., Ohara O.;  
RT "Prediction of the coding sequences of unidentified human genes. XIV.  
RT The complete sequences of 100 new cDNA clones from brain which code  
RT for large proteins in vitro.";  
RL DNA Res. 6:197-205(1999).  
RN [4]  
RN SEQUENCE OF 76-1127 FROM N.A. (ISOFORM ALPHA).  
RP TISSUE=Thyroid;  
RC MEDLINE=99367902; PubMed=10439047;  
RA Klugbauer S., Rabes H.;  
RT "The transcription coactivator ht1f and a related protein are fused  
RT to the ret receptor tyrosine kinase in childhood papillary thyroid  
RT carcinomas.";  
RL Oncogene 18:4388-4393(1999).  
CC -1- FUNCTION: Seems to act as a transcriptional repressor.  
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).  
CC -1- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=2;  
CC Name=Alpha;  
CC IsoId=Q9UPN9-1; Sequence=Displayed;  
CC Name=Beta;  
CC IsoId=Q9UPN9-2; Sequence=VSP\_005774;  
CC -1- SIMILARITY: Contains 1 RING-type zinc finger.  
CC -1- SIMILARITY: Contains 2 B box-type zinc fingers.  
CC -1- SIMILARITY: Contains 1 bromodomain.  
CC -1- SIMILARITY: Contains 1 PHD-type zinc finger.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC ENBL; AF119043; AAD17259.1; .  
CC ENBL; AF220136; AAG53509.1; .  
CC ENBL; AF220137; AAG53510.1; .  
CC ENBL; AB029036; BAA83065.1; ALT\_INT.  
CC ENBL; AJ132948; CAB55313.1; .  
CC TRANSFAC; T04946; .  
CC Genew; HGNC:16290; TRIM33.  
CC MIM; 605769; .  
CC GO; GO:0005634; C:nucleus; NAS.  
CC GO; GO:0003714; F:transcription co-repressor activity; ISS.  
CC GO; GO:0008270; F:zinc ion binding; NAS.  
CC GO; GO:0016481; P:negative regulation of transcription; NAS.  
CC InterPro; IPR003649; Bbox\_C.  
CC InterPro; IPR001487; Bromodomain.  
CC InterPro; IPR000315; Znf\_Bbox.  
CC InterPro; IPR001965; Znf\_PHD.  
CC InterPro; IPR001841; Znf\_RING.  
CC Pfam; PF00439; bromodomain; 1.  
CC Pfam; PF00628; PHD; 1.

DR Pfam; PF00643; zf-B\_box; 2.  
DR PRINTS; PRO1406; BROXZNFINGER.  
DR PRINTS; PRO0503; BROMODOMAIN.  
DR SMART; SM00502; BBC; 1.  
DR SMART; SM00336; BBOX; 2.  
DR SMART; SM00297; BROMO; 1.  
DR SMART; SM00249; PHD; 2.  
DR SMART; SM00184; RING; 2.  
DR PROSITE; PS00633; BROMODOMAIN\_1; FALSE\_NEG.  
DR PROSITE; PS00119; ZF\_BBOX; 2; 1.  
DR PROSITE; PS00119; ZF\_BBOX; 2; 1.  
DR PROSITE; PS01359; ZF\_PHD\_1; 1.  
DR PROSITE; PS00016; ZF\_PHD\_2; 1.  
DR PROSITE; PS00518; ZF\_RING\_1; 1.  
DR PROSITE; PS00089; ZF\_RING\_2; 1.  
KW Transcription regulation; Repressor; DNA-binding; Bromodomain;  
KW Zinc-finger; Nuclear protein; Repeat; Coiled coil;  
KW Alternative splicing.  
FT DOMAIN 13 18 POLY-GLY.  
FT DOMAIN 44 49 POLY-GLU.  
FT ZN\_FING 125 154 RING-TYPE.  
FT ZN\_FING 212 259 B\_BOX-TYPE 1.  
FT ZN\_FING 271 312 B\_BOX-TYPE 2.  
FT DOMAIN 299 401 COILED COIL (POTENTIAL).  
FT DOMAIN 545 550 POLY-THR.  
FT ZN\_FING 887 934 PHD-TYPE.  
FT DOMAIN 974 1046 BROMODOMAIN.  
FT VARSPLIC 1041 1057 Missing (in isoform Beta).  
FT FTIG=VSP\_005774.  
FT CONFLICT 89 89 V -> E (IN REF. 4).  
FT CONFLICT 451 453 PAA -> LLH (IN REF. 4).  
FT CONFLICT 909 909 F -> S (IN REF. 4).  
FT CONFLICT 1037 1037 R -> T (IN REF. 1).  
FT CONFLICT 1114 1127 RKELKSDERPVIK -> QKTPKVR (IN REF. 1).  
SQ SEQUENCE 1127 AA; 122521 MW; 02E3860CFDFA3B CRC64;  
Query Match 52.3%; Score 67.5; DB 1; Length 1127;  
Best Local Similarity 54.2%; Pred. No. 0.013; Mismatches 2; Indels 3; Gaps 1;  
Matches 13; Conservative  
QY 1 FRQPDVAVKGLPDYHKIKIQPWD 24  
DB 982 FQEPVPA---SIPNYKIKIKQPWD 1002  
RESULT 15  
GCN5\_YEAST STANDARD; PRT; 439 AA.  
ID GCN5\_YEAST  
AC Q03330;  
DT 01-OCT-1993 (Rel. 27, Created)  
DT 01-OCT-1993 (Rel. 27, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Histone methyltransferase GCN5 (EC 2.3.1.48).  
GN GCN5 OR ADAM OR YGR252W.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;  
RN [1]  
RN SEQUENCE FROM N.A.  
RX MEDLINE=93011009; PubMed=1396595;  
RA Georgakopoulos I., Thireos G.;  
RT "Two distinct yeast transcriptional activators require the function  
RT of the GCN5 protein to promote normal levels of transcription.";  
RL EMBO J. 11:4145-4152(1992).  
RN [2]  
RN SEQUENCE FROM N.A.  
RX MEDLINE=97313265; PubMed=9169869;  
RA Tettelin H., Agostoni Carbone M.L., Albermann K., Albers K.,  
RA Arroyo K., Backes U., Barreiros I., Bertani I., Bjorson A.J.,  
RA Brueckner M., Bruschi C.V., Carignani G., Castagnoli L., Cerdan E.,  
RA Clemente M.L., Coblenz A., Coglieva M., Coissac E., Defoor E.,  
RA Del Bino S., Delius H., Delner D., de Wergifosse P., Dujon B.,

RA Durand P., Entian K.-D., Erasó P., Escribano V., Fabiani L.,  
RA Fartmann B., Feroli F., Feuermann M., Frontali L., Garcia-Gonzalez M.,  
RA Garcia-Saez M.L., Goffeau A., Guerreiro P., Hani J., Hansen M.,  
RA Hebling U., Hernandez K., Heumann K., Hilger F., Hofmann B.,  
RA Indge K.J., James C.M., Klima R., Koetter P., Kramer B., Kramer W.,  
RA Lauquin G., Leuther H., Louis E.J., Maillier E., Marconi A.,  
RA Martegani E., Mazon M.J., Mazzoni C., McReynolds A.D.K.,  
RA Melchiorretto P., Mewes H.-W., Minenkova O., Mueller-Auer S.,  
RA Nawrocki A., Netter P., Neu R., Nombela C., Oliver S.G., Panzeri L.,  
RA Paoluzzi S., Plevani P., Portetelle D., Portillo F., Potier S.,  
RA Purnelle B., Rieger M., Riles L., Rinaldi T., Robben J.,  
RA Rodriguez-Pousada C., Rodriguez-Belmonte E., Rodriguez-Torres A.M.,  
RA Rose M., Ruzi M., Saliola M., Sanchez-Perez M., Schaefer B.,  
RA Schaefer M., Schafke M., Schmidheini T., Schreier A., Skala J.,  
RA Souciet J.-L., Steensma H.Y., Talla E., Thierry A., Vandenbol M.,  
RA van der Aart Q.J.M., Van Dyck L., Vanoni M., Verhasselt P., Voet M.,  
RA Volckaert G., Wambutt R., Watson M.D., Weber N., Wedler E., Wedler H.,  
RA Zipfl P., Wolf K., Wright L.F., Zaccaria P., Zimmermann M.,  
RA Zöllner A., Kleine K.,  
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome VII.";  
RL Nature 387:81-84(1997).  
RN [3]  
RP SEQUENCE OF 1-170 FROM N.A.  
RC STRAIN=S288C / FY1679;  
RX MEDLINE=97275234; PubMed=9133742;  
RA Feroli F., Carignani G., Pavanello A., Guerreiro P., Azevedo D.,  
RA Rodriguez-Pousada C., Melchiorretto P., Panzeri L.,  
RA Agostoni Carbone M.L.,  
RT "Analysis of a 17.9 kb region from Saccharomyces cerevisiae  
chromosome VII reveals the presence of eight open reading frames,  
including BRFL1 (TFR1B70) and GCN5 genes.";  
RL Yeast 13:373-377(1997).  
RN [4]  
RP SEQUENCE OF 170-439 FROM N.A.  
RC STRAIN=S288C;  
RX MEDLINE=97279233; PubMed=9133741;  
RA Mazzoni C., Ruzi M., Rinaldi T., Solinas F., Montebove F.,  
RA Frontali L.,  
RT "Sequence analysis of a 10.5 kb DNA fragment from the yeast  
chromosome VII reveals the presence of three new open reading frames  
and of a tRNA<sup>Thr</sup> gene.";  
RL Yeast 13:369-372(1997).  
RN [5]  
RP ASSOCIATION WITH ADA2.  
RX MEDLINE=95045371; PubMed=7957049;  
RA Marcus G.A., Silverman N., Berger S.L., Horiuchi J., Guarente L.,  
RT "Functional similarity and physical association between GCN5 and  
RT ADA2: putative transcriptional adaptors.";  
RL EMBO J. 13:4807-4815(1994).  
RN [6]  
RP IDENTIFICATION IN A SAGA COMPLEX WITH SPT2; SPT7; SPT8; SPT20; HFII  
ADA2; ADA3 AND TRAI.  
RX MEDLINE=99102959; PubMed=9885573;  
RA Grant P.A., Schieltz D., Pray-Grant M.G., Yates J.R. III,  
RA Workman J.L.,  
RT "The ATM-related cofactor Tra1 is a component of the purified SAGA  
RT complex.";  
RL Mol. Cell 2:863-867(1998).  
RN [7]  
RP REVIEW.  
RX MEDLINE=20400931; PubMed=10940244;  
RA Dyda F., Klein D.C., Hickman A.B.,  
RT "GCN5-related N-acetyltransferases: a structural overview.";  
RL Annu. Rev. Biophys. Biomol. Struct. 29:81-103(2000).  
RN [8]  
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 99-262.  
RX MEDLINE=99362688; PubMed=10430873;  
RA Triebel R.C., Rojas J.R., Steiner D.E., Venkataramani R.N., Wang L.,  
RA Zhou J., Allis C.D., Berger S.L., Marmorstein R.,  
RT "Crystal structure and mechanism of histone acetylation of the yeast  
RT GCN5 transcriptional coactivator.";  
RT Proc. Natl. Acad. Sci. U.S.A. 96:8931-8936(1999).  
CC -!- FUNCTION: Acetylates Lys-14 of histone H3. Also acetylates Lys-8

CC and Lys-16 of histone H4 with a lower preference. Acetylation of  
histones gives a specific tag for epigenetic transcription  
activation. Operates in concert with certain DNA-binding  
transcriptional activators such as GCN4 or HAP2/3/4. Acts via the  
formation of large multiprotein complexes that modify the  
chromatin.  
CC -!- CATALYTIC ACTIVITY: Acetyl-CoA + histone = CoA + acetyl-histone.  
CC -!- SUBUNIT: Interacts with ADA2. Part of the ADA/GCN5 complex that  
consists of HFII/ADA1, ADA2, ADA3, SPT20/ADA5 AND GCN5. Component  
of the SAGA complex, at least composed of SPT2, SPT7, SPT8,  
SPT20/ADA5, HFII, ADA2, ADA3/NGG1, TRAI and GCN5.  
CC -!- SUBCELLULAR LOCATION: Nuclear.  
CC -!- SIMILARITY: Contains 1 bromodomain.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; X68628; CAA48602.1; -;  
DR EMBL; Z73037; CAA97281.1; -;  
DR EMBL; X99228; CAA67614.1; -;  
DR PIR; S28051; S28051.  
DR PDB; IVGH; 02-AUG-99.  
DR PDB; 1F6I; 24-NOV-00.  
DR GeneOnline; 141564; -;  
DR SGD; S0003484; GCN5.  
DR TRANSFAC; T02145; -;  
DR GO; GO:0000124; C:SAGA complex; IDA.  
DR InterPro; IPR001487; Bromodomain.  
DR InterPro; IPR00182; GCN5acetyl\_trans.  
DR Pfam; PF00583; Acetyltrans; 1.  
DR Pfam; PF00439; bromodomain; 1.  
DR PRINTS; PR00503; BROMODOMAIN.  
DR SMART; SM00297; BROMO; 1.  
DR PROSITE; PS00633; BROMODOMAIN 1; 1.  
DR PROSITE; PS0014; BROMODOMAIN 2; 1.  
KW Transcription regulation; Transferase; Activator; Chromatin regulator;  
KW Trans-acting factor; Nuclear protein; Bromodomain; 3D-structure.  
FT ACT\_SITE 173 173  
FT DOMAIN 344 414  
FT STRAND 100 105  
FT HELIX 111 127  
FT TURN 129 130  
FT TURN 133 141  
FT TURN 143 144  
FT STRAND 146 152  
FT TURN 153 155  
FT STRAND 156 166  
FT HELIX 167 169  
FT TURN 170 170  
FT STRAND 171 179  
FT TURN 181 182  
FT TURN 186 187  
FT HELIX 188 203  
FT STRAND 208 213  
FT HELIX 215 217  
FT HELIX 218 223  
FT TURN 224 225  
FT STRAND 227 227  
FT HELIX 234 237  
FT STRAND 238 238  
FT TURN 239 240  
FT STRAND 241 241  
FT TURN 244 245  
FT STRAND 248 253  
SQ SEQUENCE 439 AA; 51069 MW; 3200730DDCTEF70D CRC64;  
Query Match 47.3%; Score 61; DB 1; Length 439;  
Best Local Similarity 54.2%; Pred. No. 0.051;

Matches 13; Conservative 4; Mismatches 5; Indels 2; Gaps 1;

Qy 1 FROPDAVKLGLPDYHKIKQPMW 24  
Db 352 FLOPVN--KEEVPDYDFIKPEMD 373

Search completed: July 27, 2004, 12:07:37  
Job time : 15 secs

Result No.	Query #			DB	ID	Description
	Score	Match	Length			
1	129	100.0	503	11	Q99PC5	Q99pc5 mus musculus
2	129	100.0	539	13	Q72VH4	Q72yh4 xenopus lae
3	129	100.0	729	13	Q03711	Q03971 gallus gall
4	129	100.0	798	11	C88411	C88411 mus musculus
5	129	100.0	798	11	C54795	C54795 mus musculus
6	117	90.7	515	13	Q78XL8	Q78xl8 brachydanio
7	113	87.6	1110	5	Q81RN6	Q81rn6 drosophila
8	113	87.6	2038	5	Q93L33	Q9w13 drosophila
9	110	85.3	505	11	C8C665	C8c665 mus musculus
10	110	85.3	511	11	C8CAX7	C8cax7 mus musculus
11	110	85.3	556	4	Q8NSM3	Q8nsm3 homo sapien
12	110	85.3	726	11	Q8KZF0	Q8kzf0 mus musculus
13	110	85.3	726	11	Q9J125	Q9j125 mus musculus
14	108	83.7	558	13	Q75XP6	Q75xp6 brachydanio
15	107	82.9	723	11	Q8VHF7	Q8vhf7 mus musculus
16	107	82.9	731	4	O60433	O60433 homo sapien

```
Query Match      100.0%; Score 129; DB 11; Length 503;
Best Local Similarity 100.0%; Pred. No. 1.4e-11;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

db 98 FROPYDAVKI.GT.PDYHKTKOPMD 121

98 FRQPVDAVKGLPDIYHKIIRKQPM 121

## RESULT 2

Q7ZYH4 PRELIMINARY; PRT; 539 AA.  
 AC Q7ZYH4  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-OCT-2002 (TrEMBLrel. 25, Last annotation update)  
 DE Similar to bromodomain-containing 2 (Fragment).  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;  
 OC Xenopodinae; Xenopus.  
 OX NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Embryo;  
 RA Klein S., Strausberg R.,  
 RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC043784; AA43784.1; -.  
 DR InterPro; IPR001487; Bromodomain.  
 DR Pfam; PF00439; bromodomain; 2.  
 DR PRINTS; PR00503; BROMODOMAIN.  
 DR SMART; SM00297; BROMO; 2.  
 DR PROSITE; PS00633; BROMODOMAIN 1; 1.  
 DR PROSITE; PS50014; BROMODOMAIN\_2; 2.  
 FT NON-TER 539  
 SQ SEQUENCE 539 AA; 60130 MW; 53841988925415F4 CRC64;

Query Match 100.0%; Score 129; DB 13; Length 539;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-11;  
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

## QY 1 FRQPVDAVKLGLPDYHKIIKQPM 24

DB 99 FRQPVDAVKLGLPDYHKIIKQPM 122

## RESULT 3

Q90971 PRELIMINARY; PRT; 729 AA.  
 AC Q90971  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE Kinase.  
 DE RING3.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bi2;  
 RA MEDLINE=56376536; PubMed=8781126;  
 RX Thorpe K.L., Abdulla S., Kaufman J., Trowsdale J., Beck S.;  
 RT "Phylogeny and structure of the RING3 gene."  
 RL Immunogenetics 44:391-396(1996).  
 DR EMBL; X96669; CAA65449.1; -.  
 DR HSP; Q92831; IPR001487; Bromodomain.  
 DR InterPro; IPR001487; Bromodomain.  
 DR Pfam; PF00439; bromodomain; 2.  
 DR PRINTS; PR00503; BROMODOMAIN.  
 DR SMART; SM00297; BROMO; 2.  
 DR PROSITE; PS00633; BROMODOMAIN 1; 2.  
 DR PROSITE; PS50014; BROMODOMAIN\_2; 2.  
 SQ SEQUENCE 729 AA; 80389 MW; 06E1B92C804DF7B8 CRC64;

Query Match 100.0%; Score 129; DB 13; Length 729;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-11;  
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FRQPVDAVKLGLPDYHKIIKQPM 24  
 DB 52 FRQPVDAVKLGLPDYHKIIKQPM 75

## RESULT 4

O88411 PRELIMINARY; PRT; 798 AA.  
 AC O88411  
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE Female sterile homeotic-related protein Frg-1.  
 GN BRD2.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Rhee K., Brunori M., Besset V., Wolgemuth D.J.;  
 RT "Expression and potential role of Frg-1, a putative murine  
 RT bromodomain-containing homologue of the Drosophila gene female sterile  
 RT homeotic."  
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF045462; AAC24810.1; -.  
 DR HSP; Q92831; IPR001487; Bromodomain.  
 DR MGB; MGI:99495; Brd2.  
 DR InterPro; IPR001487; Bromodomain.  
 DR Pfam; PF00439; bromodomain; 2.  
 DR PRINTS; PR00503; BROMODOMAIN.  
 DR SMART; SM00297; BROMO; 2.  
 DR PROSITE; PS00633; BROMODOMAIN 1; 2.  
 DR PROSITE; PS50014; BROMODOMAIN\_2; 2.  
 SQ SEQUENCE 798 AA; 88063 MW; A9942517CF15B7A1 CRC64;

Query Match 100.0%; Score 129; DB 11; Length 798;  
 Best Local Similarity 100.0%; Pred. No. 2.3e-11;  
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FRQPVDAVKLGLPDYHKIIKQPM 24  
 DB 98 FRQPVDAVKLGLPDYHKIIKQPM 121

## RESULT 5

O54795 PRELIMINARY; PRT; 798 AA.  
 AC O54795  
 DT 01-JUN-1998 (TrEMBLrel. 06, Created)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE RING3 protein.  
 GN BRD2 OR RSRG1 OR RING3.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Thorpe K.L., Beck S.;  
 RL Immunogenetics 0:0-0(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=129SVJ;  
 RA Taniguchi Y., Matsuzaka Y., Fujimoto H., Miyado K., Kohda A.,  
 RA Okumura K., Kimura M., Inoko H.;  
 RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE OF 47-549 FROM N.A.  
 RC STRAIN=INBRED CD-1; TISSUE=Testis;  
 RA Taniguchi Y., Matsuzaka Y., Fujimoto H., Miyado K., Kohda A.,  
 RA Okumura K., Kimura M., Inoko H.;







```
RP SEQUENCE FROM N.A.
RA FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003442; AAF4612.3; -.
DR HSP; Q92831; I891
DR FlyBase; FBgn0004656; fs(1)h.
DR InterPro; IPR001487; Bromodomain.
DR Pfam; PF00439; bromodomain; 2.
DR PRINTS; PR00503; BROMODOMAIN.
DR SMART; SM00297; BROMO; 2.
DR PROSITE; PS00633; BROMODOMAIN_1; 2.
DR PROSITE; PS50014; BROMODOMAIN_2; 2.
DR PROSITE; PS00583; PFKB_KINASES_1; 1.
SQ SEQUENCE 2038 AA; 205346 MW; DC4A1A7B1266191E CRC64;

Query Match      87.6%; Score 113; DB 5; Length 2038;
Best Local Similarity 87.3%; Pred. No. 21e-08;
Matches 21; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 FRQPDVAVKGLPDYHKIIKPMD 24
   ||||| ||||| ||||| |||||
Db 59 FQQPDVDAKLNLPDYHKIIKPMD 82

RESULT 9
ID Q8C665 PRELIMINARY; PRT; 505 AA.
AC Q8C665;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Bromodomain-containing 3 (Fragment).
GN BRD3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium.
RT "Analysis of the mouse transcriptome based on functional annotation of
   60,770 full-length cDNAs."
RL Nature 420:563-573 (2002).
DR EMBL; AK076472; BAC36359.1; -.
DR MGD; MGI:1914632; Brd3.
DR InterPro; IPR001487; Bromodomain.
DR Pfam; PF00439; bromodomain; 2.
DR PRINTS; PR00503; BROMODOMAIN.
DR SMART; SM00297; BROMO; 2.
DR PROSITE; PS00633; BROMODOMAIN_1; 2.
DR PROSITE; PS50014; BROMODOMAIN_2; 2.
FT NON_TER 505
SQ SEQUENCE 505 AA; 55211 MW; 35C76A91AAEC6AB4 CRC64;

Query Match      85.3%; Score 110; DB 11; Length 505;
Best Local Similarity 83.3%; Pred. No. 1.3e-08;
Matches 20; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 FRQPDVAVKGLPDYHKIIKPMD 24
   ||||| ||||| ||||| |||||
Db 58 FYQPDVDAKLNLPDYHKIIKPMD 81

RESULT 10
Q8CAX7
ID Q8CAX7 PRELIMINARY; PRT; 511 AA.
AC Q8CAX7;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
```

```
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Bromodomain-containing 3 (Fragment).
GN BRD3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium.
RT "Analysis of the mouse transcriptome based on functional annotation of
   60,770 full-length cDNAs."
RL Nature 420:563-573 (2002).
DR EMBL; AK037435; BAC29806.1; -.
DR MGD; MGI:1914632; Brd3.
DR InterPro; IPR001487; Bromodomain.
DR Pfam; PF00439; bromodomain; 2.
DR PRINTS; PR00503; BROMODOMAIN.
DR SMART; SM00297; BROMO; 2.
DR PROSITE; PS00633; BROMODOMAIN_1; 2.
DR PROSITE; PS50014; BROMODOMAIN_2; 2.
FT NON_TER 511
SQ SEQUENCE 511 AA; 55921 MW; E189A3484ABE2679 CRC64;

Query Match      85.3%; Score 110; DB 11; Length 511;
Best Local Similarity 83.3%; Pred. No. 1.3e-08;
Matches 20; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 FRQPDVAVKGLPDYHKIIKPMD 24
   ||||| ||||| ||||| |||||
Db 58 FYQPDVDAKLNLPDYHKIIKPMD 81

RESULT 11
Q8N5M3 PRELIMINARY; PRT; 556 AA.
ID Q8N5M3
AC Q8N5M3;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Similar to bromodomain containing 3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RA Strausberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC032124; AAH32124.1; -.
DR InterPro; IPR001487; Bromodomain.
DR Pfam; PF00439; bromodomain; 2.
DR PRINTS; PR00503; BROMODOMAIN.
DR SMART; SM00297; BROMO; 2.
DR PROSITE; PS00633; BROMODOMAIN_1; 2.
DR PROSITE; PS50014; BROMODOMAIN_2; 2.
SQ SEQUENCE 556 AA; 60942 MW; 8352F5DF1801A793 CRC64;

Query Match      85.3%; Score 110; DB 4; Length 556;
Best Local Similarity 83.3%; Pred. No. 1.5e-08;
Matches 20; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 FRQPDVAVKGLPDYHKIIKPMD 24
   ||||| ||||| ||||| |||||
Db 59 FYQPDVDAKLNLPDYHKIIKPMD 82

RESULT 12
Q8K2F0
```

```

ID Q8K2F0 PRELIMINARY; PRT; 726 AA.
AC Q8K2F0;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Bromodomain-containing 3.
GN BRD3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC031536; AAH31536.1; -.
DR MGD; MGI:1914632; Brd3.
DR InterPro; IPR001487; Bromodomain.
DR Pfam; PF00439; bromodomain; 2.
DR PRINTS; PR00503; bromodomain.
DR SMART; SM00297; BROMO; 2.
DR PROSITE; PS00633; BROMODOMAIN_1; 2.
DR PROSITE; PS00633; BROMODOMAIN_2; 2.
DR PROSITE; PS00633; BROMODOMAIN_2; 2.
SQ SEQUENCE 726 AA; 79791 MW; 4BC2FC0ACA41D5CD CRC64;

Query Match 85.3%; Score 110; DB 11; Length 726;
Best Local Similarity 83.3%; Pred. No. 2e-08;
Matches 20; Conservative 1; Mismatches 0; Gaps 0;

Qy 1 FROPVDAVKGLPDYHKIKQPM 24
Db 58 FYQFVDAIKLNPYHKIKQPM 81

RESULT 13
Q9J125 PRELIMINARY; PRT; 726 AA.
ID Q9J125;
AC Q9J125;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Bromodomain-containing FSH-like protein FSRG2.
GN BRD3 OR 2410084F24RIK OR FSRG2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Shang E., Wolgemuth D.J.;
RL "Cloning and expression pattern of Fshr2, a putative murine
RT bromodomain-containing homolog of the Drosophila gene female sterile
RT homeotic."
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF269193; AAF78072.1; -.
DR HSSP; Q92831; I891.
DR MGD; MGI:1914632; Brd3.
DR InterPro; IPR001487; Bromodomain.
DR Pfam; PF00439; bromodomain; 2.
DR PRINTS; PR00503; BROMODOMAIN.
DR SMART; SM00297; BROMO; 2.
DR PROSITE; PS00633; BROMODOMAIN_1; 2.
DR PROSITE; PS00633; BROMODOMAIN_2; 2.
DR PROSITE; PS00633; BROMODOMAIN_2; 2.
SQ SEQUENCE 726 AA; 79745 MW; 7AB3B4DAD38A78F4 CRC64;

Query Match 85.3%; Score 110; DB 11; Length 726;
Best Local Similarity 83.3%; Pred. No. 2e-08;
Matches 20; Conservative 1; Mismatches 0; Gaps 0;

Qy 1 FROPVDAVKGLPDYHKIKQPM 24
Db 58 FYQFVDAIKLNPYHKIKQPM 81

RESULT 14
Q7SXP6 PRELIMINARY; PRT; 558 AA.
ID Q7SXP6;
AC Q7SXP6;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein (Fragment).
OS Brachydanio rerio (zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
CX Cyprinidae; Danio.
CX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahney J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RL "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC055508; AAH55508.1; -.
KW Hypothetical protein.
FT NON TER 558 558
SQ SEQUENCE 558 AA; 62059 MW; 61681C83FBB8FA3D CRC64;

Query Match 83.7%; Score 108; DB 13; Length 558;
Best Local Similarity 79.2%; Pred. No. 3e-08;
Matches 19; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 FROPVDAVKGLPDYHKIKQPM 24
Db 73 FHEPVDATRLNLPDYHKIKQPM 96

RESULT 15
Q8VHF7 PRELIMINARY; PRT; 723 AA.
ID Q8VHF7;
AC Q8VHF7;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Bromodomain-containing protein BRD4 short variant.
GN BRD4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RA Houzelstein D., Bullock S.L., Lynch D.E., Grigorieva E.F.,
RA Wilson V.A., Beddington R.S.P.;
RT "Growth and early post implantation defects in mice mutant for the
RT bromodomain-containing protein Brd4.";
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF461396; AAL67834.1; -.
DR InterPro; IPR001487; Bromodomain.
DR Pfam; PF00439; bromodomain; 2.
DR PRINTS; PR00503; BROMODOMAIN.
DR SMART; SM00297; BROMO; 2.
DR PROSITE; PS00633; BROMODOMAIN_1; 1.
DR PROSITE; PS00014; BROMODOMAIN_2; 2.
SQ SEQUENCE 723 AA; 80635 MW; 738F2AE5F58A56BC CRC64;

Query Match      82.9%; Score 107; DB 11; Length 723;
Best Local Similarity 83.3%; Pred. No. 5.8e-08;
Matches 20; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 FRQPVDAVLGLPDYHKIIKQPMQ 24
Db 83 FQGPVDAVLNLDPYKIKITPMQ 106

Search completed: July 27, 2004, 12:08:24
Job time : 36 secs

```

